

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:44:57 ; Search time 31.9642 Seconds  
(without alignments)  
1763.712 Million cell updates/sec

Title: US-10-806-288-15  
Perfect score: 5646  
Sequence: 1 MINKNNLLTKKPIANKSN.....PAGLGALLGRKRNRKKN 1092

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/ias/5A COMB.pep:\*  
2: /cgn2\_6/ptodata/2/ias/5B COMB.pep:\*  
3: /cgn2\_6/ptodata/2/ias/6A COMB.pep:\*  
4: /cgn2\_6/ptodata/2/ias/6B COMB.pep:\*  
5: /cgn2\_6/ptodata/2/ias/PCTUS COMB.pep:\*  
6: /cgn2\_6/ptodata/2/ias/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4490	79.5	930	4	US-09-134-001C-5314
2	4485	79.4	930	4	US-09-386-962C-10
3	2698.5	47.8	1166	4	US-09-200-6508-7
4	2303.5	40.8	1742	4	US-09-386-962C-4
5	1982	35.1	930	4	US-09-200-6508-3
6	1880	33.3	1315	4	US-09-200-6508-5
7	1629.5	28.9	936	4	US-08-956-171B-5249
8	1628	28.8	933	3	US-08-933-728-2
9	1613.5	28.6	918	4	US-09-421-868-2
10	1613.5	28.6	918	4	US-09-200-6508-1
11	671.5	11.9	487	4	US-09-386-962C-14
12	635.5	11.3	2137	4	US-09-134-001C-4463
13	488	8.6	1027	4	US-08-956-171B-5254
14	482	8.5	466	4	US-09-134-001C-4749
15	408.5	7.2	669	4	US-09-107-532A-6532
16	377.5	6.7	1183	4	US-09-134-001C-3530
17	349	6.2	345	3	US-08-856-253-7
18	331.5	5.9	1161	4	US-09-327-536-2
19	315.5	5.6	1112	2	US-08-714-402-2
20	301.5	5.3	1231	3	US-08-904-263A-4
21	301.5	5.3	1231	4	US-09-434-123A-4
22	295	5.2	2315	4	US-09-543-681A-5434
23	292.5	5.2	886	4	US-08-956-171B-5235
24	292.5	5.2	2504	4	US-09-328-352-5821
25	286	5.1	1060	4	US-09-911-393-2
26	282	5.0	3052	2	US-08-557-122A-26
27	282	5.0	3052	4	US-09-262-666-26

28	279	4.9	3696	4	US-09-134-001C-5080	Sequence 5080, Ap
29	278	4.9	699	4	US-09-134-001C-4054	Sequence 4054, Ap
30	276.5	4.9	1085	1	US-08-431-080-28	Sequence 28, Appl
31	276.5	4.9	1085	2	US-08-938-534-28	Sequence 28, Appl
32	276.5	4.9	1085	4	US-09-345-294-28	Sequence 28, Appl
33	276	4.9	2391	2	US-08-446-855A-2	Sequence 2, Appl
34	276	4.9	2391	3	US-09-150-741-2	Sequence 2, Appl
35	271	4.8	1115	2	US-08-568-459A-2	Sequence 2, Appl
36	271	4.8	1115	2	US-08-487-826B-2	Sequence 2, Appl
37	271	4.8	1115	4	US-09-210-288-2	Sequence 2, Appl
38	271	4.8	1115	6	5198347-6	Patent No. 5198347
39	270.5	4.8	1833	4	US-08-621-944A-4	Sequence 4, Appl
40	270.5	4.8	1833	4	US-08-945-567D-4	Sequence 4, Appl
41	270.5	4.8	1992	4	US-08-621-944A-3	Sequence 3, Appl
42	270.5	4.8	1992	4	US-08-945-567D-3	Sequence 3, Appl
43	269.5	4.8	10182	4	US-09-134-001C-3159	Sequence 3159, Ap
44	269	4.8	60	4	US-09-386-962C-19	Sequence 19, Appl
45	267	4.7	2123	3	US-08-968-685A-10	Sequence 10, Appl

## ALIGNMENTS

### RESULT 1

US-09-134-001C-5314

; Sequence 5314, Application US/09134001C

; Patent No. 5380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; PRIOR FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 5314

; LENGTH: 930

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

; US-09-134-001C-5314

Query Match 79.5%; Score 4490; DB 4; Length 930;

Best Local Similarity 81.2%; Pred. No. 1.1e-241;

Matches 884; Conservative 19; Mismatches 26; Indels 160; Gaps 3;

QY	4	KGNLLTKKPIANKSNKYAIRKFTVTGTAIVIGATLLFGLGHNEAKAEENSVDVKDSN	63
DB	2	KGNLLTKKPIANKSNKYAIRKFTVTGTAIVIGATLLFGLGHNEAKAEENTVDVKDSN	61
QY	64	TDELSDSQSSDEKNDVINNQSDINTDDNQI IKKEETNYVDGIEKSEDTSTTN	123
DB	62	MODELSDSQSSNEEKNDVINNQSDINTDDNQ IKKEETNSNDATENRSDITQSTTN	120
QY	124	VDENEATFLQKTPQDMNTHLTERRVKSSSVESSENSISDTAQPSHTTINREESVQTSNDV	183
DB	121	VDENEATFLQKTPQDMNTHLTERRVKSSSVESSENSISDTAQPSHTTINREESVQTSNDV	180
QY	184	EDSHVSDPANSKIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQDELLN	243
DB	181	ENSRVSDPANSKIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQDELLN	240
QY	244	LPINEYENKARPLSTTSQAQPSKRVTVNQAQSGSNVHLIKVTPQSTITEGYDDSEGI	303
DB	241	LPINEYENKARPLSTTSQAQPSKRVTVNQAQSGSNVHLIKVTPQSTITEGYDDSEGI	300
QY	304	KAHDAENLIYDVFTEKVDKVGSDTMTVDIDKNTVPSDLTDSFTPIKIKNSSEIATGT	363
DB	301	KAHDAENLIYDVFTEKVDKVGSDTMTVDIDKNTVPSDLTDSFTPIKIKNSSEIATGT	360

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:49:18 ; Search time 128.509 Seconds  
(without alignments)  
2734.476 Million cell updates/sec

Title: US-10-806-288-15  
Perfect score: 5646  
Sequence: 1 MINKKNLLTKKKPIANKSN.....FAGLGALLGRRKRNKNKN 1092

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database : Published Applications AA:\*
- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
  - 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
  - 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5646	100.0	1092	15	US-10-378-674-8
2	4485	79.4	930	12	US-10-615-383-10
3	4485	79.4	930	16	US-10-689-082-10
4	4485	79.4	930	16	US-10-689-082-10
5	4309	76.3	892	16	US-10-661-809-21
6	4307	76.3	892	12	US-10-661-809-21
7	3351.5	59.4	670	12	US-10-282-122A-70481
8	2787	49.4	549	15	US-10-282-122A-70444
9	2624.5	46.5	560	15	US-10-378-674-9
10	2586	45.8	1141	12	US-10-282-122A-70251
11	2339.5	41.4	1633	12	US-10-282-122A-70437
12	2303.5	40.8	1742	12	US-10-615-383-4
13	2303.5	40.8	1742	16	US-10-689-082-4
14	2303.5	40.8	1742	16	US-10-690-184-4
15	2234	39.6	1385	12	US-10-282-122A-44324

16	2205.5	39.1	1920	12	US-10-282-122A-71413	Sequence 71413, A
17	2093.5	37.1	953	12	US-10-282-122A-44457	Sequence 44457, A
18	2069	36.6	1349	9	US-09-815-242-5898	Sequence 5898, Ap
19	2069	36.6	1349	9	US-09-815-242-13137	Sequence 13137, A
20	2011	35.6	932	9	US-09-815-242-5578	Sequence 5578, Ap
21	2011	35.6	932	9	US-09-815-242-12438	Sequence 12438, A
22	1672	29.6	343	15	US-10-378-674-4	Sequence 4, Appli
23	1634	28.9	1021	9	US-09-815-242-5471	Sequence 5471, Ap
24	1634	28.9	1021	9	US-09-815-242-12544	Sequence 12544, A
25	1629.5	28.9	936	8	US-08-781-986A-5249	Sequence 5249, Ap
26	1629.5	28.9	936	12	US-10-329-624-5249	Sequence 5249, Ap
27	1581	28.0	935	12	US-10-282-122A-44326	Sequence 44326, A
28	1562	27.7	316	15	US-10-378-674-6	Sequence 6, Appli
29	1536	27.2	877	12	US-10-282-122A-70428	Sequence 70428, A
30	1441.5	25.5	841	9	US-09-815-242-5779	Sequence 5779, Ap
31	1441.5	25.5	841	9	US-09-815-242-12751	Sequence 12751, A
32	1121.5	19.9	513	12	US-10-282-122A-59154	Sequence 59154, A
33	1023	18.1	265	13	US-10-073-256-78	Sequence 78, Appli
34	865	15.3	1113	9	US-09-815-242-5836	Sequence 5836, Ap
35	836	14.8	897	9	US-09-815-242-12769	Sequence 12769, A
36	782.5	13.9	903	12	US-10-282-122A-60847	Sequence 60847, A
37	708.5	12.5	1253	14	US-10-363-798-2	Sequence 2, Appli
38	671.5	11.9	487	12	US-10-615-383-14	Sequence 14, Appli
39	671.5	11.9	487	16	US-10-689-082-14	Sequence 14, Appli
40	671.5	11.9	487	16	US-10-690-184-14	Sequence 14, Appli
41	624.5	11.1	2271	12	US-10-282-122A-43924	Sequence 43924, A
42	615.5	10.9	2283	14	US-10-172-502-4	Sequence 4, Appli
43	607	10.8	2344	9	US-09-815-242-12713	Sequence 12713, A
44	594	10.5	1831	12	US-10-282-122A-71033	Sequence 71033, A
45	554	9.8	406	16	US-10-451-467A-262	Sequence 262, App

ALIGNMENTS

RESULT 1

US-10-378-674-8  
; Sequence 8, Application US/10378674  
; Publication No. US20040006209A1  
; GENERAL INFORMATION:  
; APPLICANT: PANTI Joseph M.  
; TITLE OF INVENTION: MONOCLONAL AND POLYCLONAL ANTIBODIES RECOGNIZING COAGULASE-NEGATIV  
; FILE REFERENCE: P07556US01/BAS  
; CURRENT APPLICATION NUMBER: US/10/378,674  
; CURRENT FILING DATE: 2003-03-05  
; PRIOR APPLICATION NUMBER: 60/361,324  
; PRIOR FILING DATE: 2002-03-05  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 1092  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-10-378-674-8

Query Match	100.0%	Score	5646	DB	15	Length	1092
Best Local Similarity	100.0%	Pred	No. 3.8e-307				
Matches	1092	Conservative	0	Mismatches	0	Indels	0
						Gaps	0
QY	1	MINKKNLLTKKKPIANKSNKYAIRKFTVTGTAIVIGATLLFGLGHNEAKAENSVDVK	60				
Db	1	MINKKNLLTKKKPIANKSNKYAIRKFTVTGTAIVIGATLLFGLGHNEAKAENSVDVK	60				
QY	61	DSNTDELSDSDNDSDDEKNDVNNQSDNTDNNQIIKKETNNYDGIKESDRTES	120				
Db	61	DSNTDELSDSDNDSDDEKNDVNNQSDNTDNNQIIKKETNNYDGIKESDRTES	120				
QY	121	TTNVDENATFLQKTPQDNTHLTTEBEVKSSSVSSNSSIDTAQPSHTTINREESVQTS	180				
Db	121	TTNVDENATFLQKTPQDNTHLTTEBEVKSSSVSSNSSIDTAQPSHTTINREESVQTS	180				
QY	181	DNVEDSHVSDPANSKIKESNTSGKEENTIEQPNKVEDSTTSQPSGYTNIDEKISNQDE	240				

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:42:37 ; Search time 39.7921 Seconds  
(without alignments)  
2639.748 Million cell updates/sec

Title: US-10-806-288-15

Perfect score: 5646

Sequence: 1 MINKNNLLTKKPIANKSN.....PAGLGALLGKRRKRNKN 1092

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1.\*  
2: Pir2.\*  
3: Pir3.\*  
4: Pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5646	100.0	1092	2 T30214	fibrinogen-binding
2	2698.5	47.8	1166	2 T28680	fibrinogen-binding
3	2589	45.9	1141	2 D89824	hypothetical prote
4	2234	39.6	1385	2 D89824	hypothetical prote
5	2093.5	37.1	953	2 C89824	hypothetical prote
6	1880	33.3	1315	2 T28679	fibrinogen-binding
7	1652	29.3	989	2 D89852	fibrinogen-binding
8	1628	28.8	933	2 S41539	fibrinogen-binding
9	1536	27.2	877	2 F90070	Clumping factor B
10	785	13.9	882	2 AG1671	probable peptidogl
11	782.5	13.9	903	2 AG1299	probable peptidogl
12	624.5	11.1	2271	2 F90073	hypothetical prote
13	554	9.8	406	2 S38170	SRP40 protein - ye
14	554	9.8	3394	2 T18501	hypothetical prote
15	549	9.7	2570	2 T17451	fimbriae-associate
16	525.5	9.3	1038	2 T90053	hypothetical prote
17	520	9.2	940	2 S19702	fibronectin-bindin
18	502.5	8.9	961	2 G90053	hypothetical prote
19	487.5	8.6	1018	2 A32192	fibronectin-bindin
20	471	8.3	334	2 A54138	acidic repetitive
21	471	8.3	640	2 A54502	S antigen precurs
22	452	8.0	2910	2 T28156	DNA-directed RNA p
23	442.5	7.8	1999	2 AB2018	hypothetical prote
24	435	7.7	1192	2 A71623	probable secreted
25	430.5	7.6	4550	2 T18440	hypothetical prote
26	391.5	6.9	1217	2 S52714	sericinB - silkwo
27	388.5	6.9	695	2 S27390	calcium-binding pr
28	381.5	6.8	955	2 T18435	hypothetical prote
29	379.5	6.7	3724	2 T18427	hypothetical prote

#### RESULT 1

T30214  
fibrinogen-binding protein - Staphylococcus epidermidis  
C:Species: Staphylococcus epidermidis  
C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 11-May-2000  
C:Accession: T30214  
R:Nilsson, M.; Frykberg, L.; Flock, J.I.; Pei, L.; Lindberg, M.; Guss, B.  
Infect. Immun. 66, 2666-2673, 1998  
A:Title: A Fibrinogen-binding protein of Staphylococcus epidermidis.  
A:Reference number: Z20781; MUID:98261511; PMID:9596732  
A:Accession: T30214  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1092 <NIL>  
A:Cross-references: EMBL:Y17116; NID:e1296734; PID:e1296735; PIDN:CAA76638.1

Query Match 100.0%; Score 5646; DB 2; Length 1092;

Best Local Similarity 100.0%; Pred. No. 3.3e-211;

Matches 1092; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MINKNNLLTKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDVK	60
Db	1	MINKNNLLTKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDVK	60
Qy	61	DSNTDDELSDNQSSDEERKNDVINNNQSIINTDNNQIIKKBETNNYDGIKRSSEDRTES	120
Db	61	DSNTDDELSDNQSSDEERKNDVINNNQSIINTDNNQIIKKBETNNYDGIKRSSEDRTES	120
Qy	121	TTNVDENEATFLOKTPQDNTHLTREEVKSSSVSSSSSIDTAQOPSHTTTINREESVQTS	180
Db	121	TTNVDENEATFLOKTPQDNTHLTREEVKSSSVSSSSSIDTAQOPSHTTTINREESVQTS	180
Qy	181	DNVEDSHVSPFANSKIKESNTSGKENTIEOPKNKVEDSTTSOPSGVTNIDKISNODE	240
Db	181	DNVEDSHVSPFANSKIKESNTSGKENTIEOPKNKVEDSTTSOPSGVTNIDKISNODE	240
Qy	241	LLNLPINEYENKARPLSTTSAQPSIKKVTNVQLAAEQSNVNHIIKVTQDSITEGYDDSE	300
Db	241	LLNLPINEYENKARPLSTTSAQPSIKKVTNVQLAAEQSNVNHIIKVTQDSITEGYDDSE	300
Qy	301	GVTKAHDAENLIYDVTFPEVDKVKSGDTMTVDIDKNTVPSDLTSPFTPKIKNSGRIIA	360
Db	301	GVTKAHDAENLIYDVTFPEVDKVKSGDTMTVDIDKNTVPSDLTSPFTPKIKNSGRIIA	360
Qy	361	TGTYDKNKQITTFPTDYVDKYNIIKAHLKLTSTVIDSKVPNNNTKLDVETKALSSVYK	420
Db	361	TGTYDKNKQITTFPTDYVDKYNIIKAHLKLTSTVIDSKVPNNNTKLDVETKALSSVYK	420
Qy	421	TTTVEYQRPNNENTANLQSNFTNIDTKNHTVEQTIYINPLRYSAKETNNVNSGNGDEGT	480
Db	421	TTTVEYQRPNNENTANLQSNFTNIDTKNHTVEQTIYINPLRYSAKETNNVNSGNGDEGT	480

#### ALIGNMENTS

30	371.5	6.6	1305	2	T00670
31	370	6.6	1063	2	D86731
32	368	6.5	1664	2	T18262
33	366	6.5	3844	2	T18402
34	364.5	6.5	792	2	T42963
35	363.5	6.4	1360	2	T18403
36	362	6.4	1193	2	G71605
37	351.5	6.2	1428	2	T08852
38	345.5	6.1	2481	2	D90011
39	343.5	6.1	1658	2	S55101
40	338	6.0	4776	2	E95206
41	336	6.0	2485	1	H71621
42	335.5	5.9	4910	2	S64942
43	333.5	5.9	1072	2	A86827
44	333	5.9	797	2	A36811
45	331.5	5.9	1419	2	T30531

probable inositol  
hypothetical prote  
S-layer protein -  
asparagine/asparta  
hypothetical prote  
asparagine/asparta  
hypothetical prote  
lustrin A - Califo  
FmB protein (limpo  
hypothetical prote  
cell wall surface  
serine/threonine-s  
probable membrane  
hypothetical prote  
hypothetical prote  
agglutinin-like ad

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:34:15 ; Search time 22.8315 Seconds  
(without alignments)  
2490.441 Million cell updates/sec

Title: US-10-806-288-15  
Perfect score: 5646  
Sequence: 1 MINKNNLLTKKPIANKSN.....FAGLGALLGKKRKNRKN 1092

Scoring table: BLOSUM62

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1319.5	23.4	1637	1 MRSP STAAU	P80544 staphylococ
2	708.5	12.5	1253	1 DSPP HUMAN	Q9NZM4 homo sapien
3	675.5	12.0	934	1 DSPP MOUSE	P97399 mus musculu
4	606.5	10.7	687	1 DSPP RAT	Q62598 rattus norv
5	554	9.8	406	1 SR40 YEAST	P32583 saccharomyc
6	487.5	8.6	1018	1 FNBA STAAU	P14738 staphylococ
7	471	8.3	640	1 SNT PLAFW	P13821 plasmodium
8	388.5	6.9	695	1 ARP EUGR	Q04732 euglena gra
9	368	6.5	1664	1 SLPI CLOTM	Q06852 dictyosteli
10	361.5	6.4	400	1 RTOA DICDI	Q03661 saccharomyc
11	343.5	6.1	1658	1 YM67 YEAST	Q12019 saccharomyc
12	335.5	5.9	4910	1 MDN1 YEAST	Q01033 herpesvirus
13	333	5.9	797	1 VG48 HSVSA	Q13368 candida alb
14	331.5	5.9	1419	1 ALAI CANAL	P46591 candida alb
15	315	5.6	937	1 HYRI CANAL	Q7YQM4 pan troglod
16	314	5.6	2492	1 ATRX PANTR	P14248 plasmodium
17	311	5.5	2452	1 RPBI PLAFD	P47035 saccharomyc
18	310	5.5	1189	1 YJH6 YEAST	P46100 homo sapien
19	303	5.4	2492	1 ATRX HUMAN	O55188 mus musculu
20	301.5	5.3	503	1 DMPI MOUSE	Q7YQM3 pongo pygma
21	297	5.3	2492	1 ATRX PONPY	P98193 rattus norv
22	295.5	5.2	489	1 DMPI RAT	P18899 saccharomyc
23	294.5	5.2	429	1 DR48 YEAST	Q01550 xenopus lae
24	291	5.2	1744	1 TANA XENLA	P07856 bombyx mori
25	288	5.1	389	1 SERI BOMMO	Q13316 homo sapien
26	285.5	5.1	513	1 DMPI HUMAN	P39653 streptococc
27	283.5	5.0	1337	1 DEXT STRDO	P39793 bacillus su
28	283	5.0	514	1 FBPA BACSU	P39520 saccharomyc
29	276.5	4.9	1085	1 IFH1 YEAST	P41891 schizosacch
30	276	4.9	500	1 GAR2 SCHPO	Q12176 saccharomyc
31	271.5	4.8	1025	1 MK21 YEAST	Q02510 staphylococ
32	271	4.8	688	1 LIP STAPB	Q61687 mus musculu
33	271	4.8	2476	1 ATRX MOUSE	

## ALIGNMENTS

### RESULT 1

MRSP STAAU STANDARD; PRT; 1637 AA  
AC P80544; O92F62;  
DT 01-FEB-1995 (Rel. 33, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Methicillin-resistant surface protein precursor.  
GN PLS.  
OS Staphylococcus aureus.  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=1280;  
RN [1]  
RC SEQUENCE FROM N.A.  
RP STRAIN=Isolate 1061;  
RX MEDLINE=21189215; PubMed=11292719;  
RA Savolainen K., Paulin L., Westerlund-Wikstrom B., Foster T.J.,  
RA Korhonen T.K., Kuusela P.;  
RT "Expression of *pis*, a gene closely associated with the *mecA* gene of  
RT methicillin-resistant *Staphylococcus aureus*, prevents bacterial  
RT adhesion in vitro."  
RL Infect. Immun. 69:3013-3020(2001).  
RN [2]  
RP SEQUENCE OF 624-628; 676-682; 938-948; 1156-1168; 1176-1185;  
RP 1199-1205 AND 1217-1224.  
RC STRAIN=Isolate 1061;  
RX MEDLINE=96270743; PubMed=8665912;  
RA Hilden P., Savolainen K., Tynnelae J., Vuento M., Kuusela P.;  
RT "Purification and characterisation of a plasmin-sensitive surface  
RT protein of *Staphylococcus aureus*."  
RL Eur. J. Biochem. 236:904-910(1996).  
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by  
CC an amide bond (Potential).

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; AF115379; AAD09131.1; -  
DR InterPro; IPR005877; Gpos\_YSIIRK.  
DR InterPro; IPR001899; Gram\_pos\_anchor.  
DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
DR Pfam; PF04650; YSIIRK\_signal; 1.  
DR TIGRFAMs; TIGR01167; LPXTG\_anchor; 1.  
DR TIGRFAMs; TIGR01168; YSIIRK\_signal; 1.  
DR PROSITE; PS00847; GRAM\_POS\_ANCHORING; 1.  
KW Antibiotic resistance; Glycoprotein; Cell wall; Peptidoglycan-anchor;  
Repeat; Signal.  
FT SIGNAL 1 48 POTENTIAL.  
FT CHAIN 49 1601 METHICILLIN-RESISTANT SURFACE PROTEIN.  
FT PROPEP 1502 1637 REMOVED BY SORTASE (POTENTIAL).  
FT DOMAIN 1301 1582 141 X 2 AA TANDEM REPEATS OF D-[SAG].

Q09624 caenorhabdi  
P22290 plasmodium  
Q00799 plasmodium  
P87498 gallus gall  
Q03400 plasmodium  
P35662 bos taurus  
P34487 caenorhabdi  
O84419 chlamydia t  
P08723 rattus norv  
P46949 saccharomyc  
Q95120 bos taurus

34 269.5 4.8 3178 1 YS89 CAEEL  
35 267.5 4.7 1070 1 PVDR PLAVB  
36 267 4.7 2867 1 RBP2 PLAVB  
37 266.5 4.7 1912 1 VITI\_CHICK  
38 265 4.7 593 1 SNT PLAF7  
39 264.5 4.7 667 1 CYLI BOVIN  
40 264.5 4.7 936 1 FHL1 YEAST  
41 264 4.7 918 1 YMJ8 CAEEL  
42 263.5 4.7 1770 1 PMPC CHLTR  
43 262.5 4.6 279 1 SPBP RAT  
44 262.5 4.6 817 1 YG4A YEAST  
45 262 4.6 510 1 DMPI BOVIN



GanCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:41:22 ; Search time 122.638 Seconds  
(without alignments)  
2809.455 Million cell updates/sec

Title: US-10-806-288-15  
Perfect score: 5646  
Sequence: 1 MINKNNLLTKKPIANKSN.....FAGLGALLGKRKRNRKN 1092

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phage:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_rvirus:\*

16: sp\_bacteriap:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5646	100.0	1092	2	070022
2	5440	96.4	1056	16	Q8CQ72
3	4485	79.4	931	2	Q9K113
4	2756	48.8	1171	2	Q9KWX6
5	2698.5	47.8	1166	2	Q86499
6	2589	45.9	1141	16	Q99W46
7	2589	45.9	1141	16	Q8NXX5
8	2586	45.8	1141	16	Q932P7
9	2344.5	41.5	1633	16	Q8CMP4
10	2303.5	40.8	1733	2	Q9K114
11	2234	39.6	1385	16	Q99W47
12	2093.5	37.1	953	16	Q99W48
13	2075.5	36.8	955	16	Q8NXX7
14	2037	36.1	1347	16	Q8NXX6
15	2023.5	35.8	947	2	Q86487
16	2006	35.5	1893	2	Q8KWL1

17	1880	33.3	1315	2	Q86488
18	1652	29.3	989	16	Q99VJ4
19	1646.5	29.2	946	16	Q8NXX1
20	1628	28.8	933	2	Q53653
21	1610	28.5	913	2	Q86476
22	1581	28.0	935	16	Q932C5
23	1573	27.9	907	16	Q8NDL0
24	1536	27.2	877	16	Q99R07
25	1471	26.1	881	2	Q93MH7
26	1216	21.5	1698	2	Q9LC00
27	1203.5	21.3	3360	16	Q88XB6
28	785	13.9	882	16	Q92AK9
29	782.5	13.9	903	16	Q8Y697
30	747.5	13.2	2020	5	Q81D80
31	671.5	11.9	487	2	Q9K112
32	670.5	11.9	970	11	Q8VBY1
33	654.5	11.6	540	2	Q8KR21
34	645.5	11.4	2310	16	Q8CMU7
35	640.5	11.3	485	16	Q8CNU7
36	639	11.3	579	13	Q77138
37	624.5	11.1	2271	16	Q99QY4
38	622	11.0	968	5	Q81SW9
39	616.5	10.9	2275	16	Q8VQJ3
40	612	10.8	2283	2	Q8VQ99
41	607	10.8	2900	5	Q812E0
42	565	10.0	1182	5	Q81D30
43	560	9.9	970	16	Q8DYL7
44	554	9.8	3394	5	Q77384
45	549	9.7	2570	2	Q9ZPF9

## ALIGNMENTS

### RESULT 1

070022 ID 070022 PRELIMINARY; PRT: 1092 AA.

AC 070022; 01-AUG-1998 (TrEMBLrel. 07, Created)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Fibrinogen-binding protein precursor.  
 OS Staphylococcus epidermidis.  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OX NCBI\_TaxID=1282;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HB;  
 RX MEDLINE=98261511; PubMed=9596732;  
 RA Nilsson M., Frykberg L., Flock J.I., Pei L., Lindberg M., Guss B.;  
 RT "A Fibrinogen-binding protein of Staphylococcus epidermidis."  
 RL Infect. Immun. 66:2666-2673(1998).  
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY AN AMIDE BOND (BY SIMILARITY).  
 DR EMBL; Y17116; CAA76638.1; -;  
 DR PIR; T30214; T30214.  
 DR GO; GO:0005618; C:cell wall; IEA.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR008454; Cna\_B.  
 DR InterPro; IPR005877; Gpos\_Ysirk.  
 DR InterPro; IPR001899; Gram\_pos\_anchor.  
 DR Pfam; PF05738; Cna\_B; 2.  
 DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
 DR Pfam; PF04650; Ysirk\_signal; 1.  
 DR TIGRFAMs; TIGR01167; LPXTG\_anchor; 1.  
 DR TIGRFAMs; TIGR01168; Ysirk\_signal; 1.  
 DR PROSITE; PS00847; GRAM\_POS\_ANCHORING; 1.  
 DR Cell wall; Peptidoglycan-anchor; Signal.  
 FT SIGNAL 1 51 POTENTIAL.  
 FT CHAIN 52 1092 FIBRINOGEN-BINDING PROTEIN.  
 SQ SEQUENCE 1092 AA; 119292 MW; 6542BC39AD8B984 CRC64;

Query Match

100.0%; Score 5646; DB 2; Length 1092;

Result No.	Query No.	Score	%		Length	DB	ID	Description
			Match					
1	1	5646	100.0		1092	2	AAW41602	Staphyloc
2	2	5646	100.0		1092	7	ABM79019	Staphyloc
3	3	4490	79.5		930	5	ABP40469	Staphyloc
4	4	4485	79.4		991	3	AAW83171	Cell wall
5	5	4485	79.4		991	3	AAW70120	Staph. ep
6	6	4307	76.3		892	6	ABU42557	Protein e
7	7	3351.5	59.4		670	6	ABU42520	Protein e
8	8	2787	49.4		549	7	ABM79020	Staphyloc
9	9	2698.5	47.8		1166	2	AAW08643	S. aureus
10	10	2696.5	47.8		1166	6	ABU18982	Pathogen
11	11	2624.5	46.5		560	7	ABM79015	Staphyloc
12	12	2586	45.8		1141	6	ABU42327	Protein e
13	13	2339.5	41.4		1633	6	ABU42513	Protein e
14	14	2303.5	40.8		1802	3	AAW83170	Cell wall
15	15	2303.5	40.8		1802	3	AAW70119	Staph. ep
16	16	2256.5	40.0		995	6	ABM72437	Staphyloc
17	17	2234	39.6		1385	6	ABU16400	Protein e
18	18	2205.5	39.1		1920	6	ABU43489	Protein e
19	19	2093.5	37.1		953	6	ABU16533	Protein e
20	20	2070	36.7		1349	6	ABM72436	Staphyloc
21	21	2069	36.6		1349	4	AAU37544	Staphyloc
22	22	2069	36.6		1349	4	AAU34402	Staphyloc
23	23	2032.5	36.0		947	6	ABU18940	Pathogen
24	24	2011	35.6		932	4	AAU36845	Staphyloc
25	25	2011	35.6		932	4	AAU34082	Staphyloc

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:41:22 ; Search time 65.362 Seconds  
(without alignments)  
2809.455 Million cell updates/sec

Title: US-10-806-288-13

Perfect score: 3013

Sequence: 1 SSDDEKNDVINNQSIINTDD.....PDGTSKSVRTDEDGKYQFDG 582

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3013	100.0	1056	16 Q8CQ72	Q8CQ72 staphylococ
2	3013	100.0	1092	2 Q70022	Q70022 staphylococ
3	2810.5	93.3	931	2 Q9K113	Q9K113 staphylococ
4	1160.5	38.5	1171	2 Q9KX16	Q9KX16 staphylococ
5	1144.5	38.0	1166	2 Q86489	Q86489 staphylococ
6	1143.5	38.0	1141	16 Q8NXX5	Q8NXX5 staphylococ
7	1137.5	37.8	1141	16 Q99W46	Q99W46 staphylococ
8	1137.5	37.8	1141	16 Q932F7	Q932F7 staphylococ
9	576	19.1	1315	2 Q86488	Q86488 staphylococ
10	567	18.8	1347	16 Q8NXX6	Q8NXX6 staphylococ
11	559	18.6	1385	16 Q99W47	Q99W47 staphylococ
12	499	16.6	1733	2 Q9K114	Q9K114 staphylococ
13	497	16.5	1633	16 Q8CMP4	Q8CMP4 staphylococ
14	489	16.2	953	16 Q99W48	Q99W48 staphylococ
15	485.5	16.1	1893	2 Q8KWM1	Q8KWM1 staphylococ
16	470.5	15.6	946	16 Q8NXXJ1	Q8NXXJ1 staphylococ

17	465	15.4	955	16 Q8NXX7	Q8NXX7 staphylococ
18	460.5	15.3	881	2 Q93MH7	Q93MH7 staphylococ
19	458.5	15.2	566	2 Q8KR22	Q8KR22 staphylococ
20	455	15.1	947	2 Q86487	Q86487 staphylococ
21	452	15.0	933	2 Q53653	Q53653 staphylococ
22	452	15.0	935	16 Q932C5	Q932C5 staphylococ
23	452	15.0	989	16 Q99V04	Q99V04 staphylococ
24	397.5	13.2	877	16 Q99R07	Q99R07 staphylococ
25	396	13.1	913	2 Q86476	Q86476 staphylococ
26	384.5	12.8	907	16 Q8NUL0	Q8NUL0 staphylococ
27	376	12.5	940	2 Q53682	Q53682 staphylococ
28	374.5	12.4	943	16 Q8NUT8	Q8NUT8 staphylococ
29	344.5	11.4	961	16 Q99RD3	Q99RD3 staphylococ
30	326.5	10.8	1015	16 Q8NUT7	Q8NUT7 staphylococ
31	326	10.8	1038	16 Q99RD2	Q99RD2 staphylococ
32	293.5	9.7	970	16 Q8DY17	Q8DY17 streptococ
33	293.5	9.7	1310	16 Q8E473	Q8E473 streptococ
34	276	9.2	2616	5 Q8IIG1	Q8IIG1 plasmodium
35	268.5	8.9	1301	5 Q8WSK5	Q8WSK5 plasmodium
36	266.5	8.8	1160	2 Q8RM86	Q8RM86 streptococ
37	266.5	8.8	1301	5 Q8IHQ2	Q8IHQ2 plasmodium
38	266	8.8	2402	2 Q9AER7	Q9AER7 staphylococ
39	265	8.8	3452	5 Q8IEA3	Q8IEA3 plasmodium
40	261	8.7	1161	2 Q9X3M7	Q9X3M7 streptococ
41	261	8.7	2153	5 Q8ICM9	Q8ICM9 plasmodium
42	259	8.6	1161	2 Q8RJ10	Q8RJ10 streptococ
43	253.5	8.4	3078	5 Q8IKH4	Q8IKH4 plasmodium
44	253.5	8.4	3207	5 Q8IHQ0	Q8IHQ0 plasmodium
45	249.5	8.3	1859	5 Q8IC27	Q8IC27 plasmodium

#### ALIGNMENTS

#### RESULT 1

Q8CQ72 PRELIMINARY; PRT; 1056 AA.  
AC Q8CQ72  
DT 01-MAR-2003 (TREMREL. 23, Created)  
DT 01-MAR-2003 (TREMREL. 23, Last sequence update)  
DT 01-OCT-2003 (TREMREL. 25, Last annotation update)  
DE Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding protein.  
DE SE0331.  
OS Staphylococcus epidermidis.  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=1282;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 12228;  
RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z., Chen Z., Wen Y.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE016745; AAC03928.1; -;  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR008454; Cna B.  
DR InterPro; IPR001899; Gram\_pos\_anchor.  
DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
DR Pfam; PF04650; YSIRK\_signal; 1.  
DR TIGRfam; TIGR01167; LPXTG\_anchor; 1.  
DR TIGRfam; TIGR01168; YSIRK\_signal; 1.  
DR PROSITE; PS00847; GRAM\_POS\_ANCHORING; 1.  
KW Complete proteome.  
SQ SEQUENCE 1056 AA; 115728 MW; 015869A9E5CA2723 CRC64;

Query Match 100.0%; Score 3013; DB 16; Length 1056;  
Best Local Similarity 100.0%; Pred.No. 2.5e-108;  
Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSDDEKNDVINNQSIINTDDNNQIIKKETHNYDIEKSEDRSTTTVDNEAFLOK 60  
|||||

Result No.	Query	Score	Match	Length	DB	ID	Description
1	300.5	10.0	1018	1	FNBA_STAAU	P14738	staphylococ
2	203	6.7	537	1	ARP_PLAPA	P04931	plasmodium
3	202	6.7	1093	1	PI4K_DICDI	P54677	dictyosteli
4	196.5	6.5	1139	1	HMW1_MYCBE	Q49413	mycoplasma
5	192.5	6.4	2867	1	RB2_PLAYB	Q00799	plasmodium
6	192	6.4	918	1	YJMB_CAEEL	P34487	caenorhabdi
7	189.5	6.3	1664	1	INT1_CANAL	P53705	candida alb
8	188.5	6.3	578	1	LIPA_MYCPU	Q50274	mycoplasma
9	188.5	6.3	1435	1	EBR1_PLAFAC	P19214	plasmodium
10	186.5	6.2	3110	1	LNMA2_HUMAN	P24043	homo sapien
11	184.5	6.1	1744	1	TANA_XENLA	Q01550	xenopus lae
12	184	6.1	817	1	YG4A_YEAST	P46949	saccharomyc
13	183.5	6.1	2009	1	SEC7_YEAST	P11075	saccharomyc
14	183	6.1	2452	1	RPB1_PLAFD	P14248	plasmodium
15	182.5	6.1	1096	1	MRC1_YEAST	P25588	saccharomyc
16	182.5	6.1	1381	1	YB87_YEAST	P34216	saccharomyc
17	182.5	6.1	1658	1	YB67_YEAST	Q03661	saccharomyc
18	182	6.0	1419	1	ALAI_CANAL	Q13368	candida alb
19	182	6.0	1460	1	N159_YEAST	P40477	saccharomyc
20	181.5	6.0	1093	1	YKD5_CAEEL	Q03563	caenorhabdi
21	179.5	6.0	0165	1	YNF4_YEAST	P53950	saccharomyc
22	179	5.9	1253	1	D3PP_HUMAN	Q9NZW4	homo sapien
23	178	5.9	960	1	YXK6_YEAST	Q04279	saccharomyc
24	178	5.9	1875	1	MLP1_YEAST	Q02455	saccharomyc
25	177	5.9	2660	1	YEEJ_ECO57	Q9X8V7	escherichia
26	176	5.8	1196	1	BXCN_CLOBO	P46081	clostridium
27	175.5	5.8	1183	1	CNA_STAAU	Q53654	staphylococ
28	175.5	5.8	1637	1	MESP_STAAU	P80544	staphylococ
29	175	5.8	1147	1	CGA1_HELPY	P80200	helicobacte
30	174	5.8	2022	1	ANT1_ONCVO	P31249	onchocerca
31	173.5	5.8	967	1	YSGO_CLOAB	P33747	clostridium
32	173	5.7	1310	1	YB35_SCHPO	Q14340	schizosacch
33	172	5.7	682	1	NISP_LACIA	Q07596	lactococcus

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:42:37 ; Search time 21.2079 Seconds  
(without alignments)  
2639.748 Million cell updates/sec

Title: US-10-806-288-13  
Perfect score: 3013  
Sequence: 1 SSDEKNDVNNQSIINTDD.....PDGTSKSVRTDEGKVFQDG 582

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues  
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: Pir1:\*  
2: Pir2:\*  
3: Pir3:\*  
4: Pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3013	100.0	1092	2 T30214	fibrinogen-binding
2	1144.5	38.0	1166	2 T28680	fibrinogen-binding
3	1137.5	37.8	1141	2 E89824	hypothetical prote
4	576	19.1	1315	2 T28679	fibrinogen-binding
5	559	18.6	1385	2 D89824	hypothetical prote
6	489	16.2	953	2 C89824	hypothetical prote
7	452	15.0	933	2 S41539	fibrinogen-binding
8	452	15.0	989	2 D89852	fibrinogen-binding
9	397.5	13.2	877	2 F90070	Clumping factor B
10	376	12.5	940	2 S19702	fibrinectin-bindin
11	344.5	11.4	961	2 G90053	hypothetical prote
12	326	10.8	1038	2 H90053	hypothetical prote
13	300.5	10.0	1018	2 A32192	fibrinectin-bindin
14	239	7.9	1039	2 T30856	protein F2 - Strept
15	225	7.5	1463	2 T30290	AAS surface protei
16	216.5	7.2	1959	2 AG1085	hypothetical prote
17	216.5	7.2	3394	2 T18501	hypothetical prote
18	212.5	7.1	1125	2 G90598	membrane nuclease,
19	211.5	7.0	2206	2 G71611	hypothetical prote
20	208.5	6.9	1072	2 A86827	hypothetical prote
21	208	6.9	1127	2 T28317	ORF MSV156 byothe
22	208	6.9	3724	2 T18427	hypothetical prote
23	207.5	6.9	1711	2 T18429	hypothetical prote
24	207.5	6.9	2401	2 T28676	rhoptry protein -
25	207	6.9	769	2 F89870	serine proteinase
26	207	6.9	1939	2 T18372	repeat organellar
27	207	6.9	4550	2 T18440	hypothetical prote
28	206.5	6.9	821	2 S67087	hypothetical prote
29	206	6.8	1302	1 JC6009	surface-located me

RESULT 1

T30214  
fibrinogen-binding protein - Staphylococcus epidermidis  
C:Species: Staphylococcus epidermidis  
C>Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 11-May-2000  
C/Accession: T30214  
R/Nilsson, M.; Frykberg, L.; Flock, J.I.; Pei, L.; Lindberg, M.; Guss, B.  
Infect. Immun. 66, 2666-2673, 1998  
A/Title: A Fibrinogen-binding protein of Staphylococcus epidermidis.  
A/Reference number: Z20781; MUID:98261511; PMID:9596732  
A/Accession: T30214  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-1092 <NIL>  
A/Cross-references: EMBL:Y17116; NID:e1296734; PID:e1296735; PIDN:CAA76638.1

Query Match 100.0%; Score 3013; DB 2; Length 1092;  
Best Local Similarity 100.0%; Pred. No. 7.2e-123;  
Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	SSDEKNDVNNQSIINTDDNNQIIKKEETNNVDGIEKESBDRTESTTNVDENAEFLQK	60
DB	75	SSDEKNDVNNQSIINTDDNNQIIKKEETNNVDGIEKESBDRTESTTNVDENAEFLQK	134
QY	61	TPQDNTHLTTEEEVKSSSVSSNSSIDTAQPSHTTINREESVQTSNDVEDSHVSDPANS	120
DB	135	TPQDNTHLTTEEEVKSSSVSSNSSIDTAQPSHTTINREESVQTSNDVEDSHVSDPANS	194
QY	121	KIKESNTSGKEENTIEQPNKVEDSTTSQPSGYTNIDEKISNODELLMLPINEYENKAR	180
DB	195	KIKESNTSGKEENTIEQPNKVEDSTTSQPSGYTNIDEKISNODELLMLPINEYENKAR	254
QY	181	PLSTTSAQPSIKRVTVNQLAARQGSNNVHLIKVTQSIITEGYDDSGVTKAHDAENLIYD	240
DB	255	PLSTTSAQPSIKRVTVNQLAARQGSNNVHLIKVTQSIITEGYDDSGVTKAHDAENLIYD	314
QY	241	VTFFVDDKVKSGDVTMTVDIDKNTVPDSDLTDSFTIPIKIDNSGEIATGYTDNKNKQIYTT	300
DB	315	VTFFVDDKVKSGDVTMTVDIDKNTVPDSDLTDSFTIPIKIDNSGEIATGYTDNKNKQIYTT	374
QY	301	FDYVDKVENIKAHKLKTSYIDKSKVPNNNTKLDVEYKTKALSSVNKTIIVTEVQRPENRT	360
DB	375	FDYVDKVENIKAHKLKTSYIDKSKVPNNNTKLDVEYKTKALSSVNKTIIVTEVQRPENRT	434
QY	361	ANLQSMFTNIDTKNHTVEQTYINFLRYSAKETNNISGNGEGSTIIDDSITIIKYKVG	420
DB	435	ANLQSMFTNIDTKNHTVEQTYINFLRYSAKETNNISGNGEGSTIIDDSITIIKYKVG	494
QY	421	DNQNLPSDNRIYDYSEYEDVTNDYQAQGNNDVNFQGNIDSPYIIKISKYDPNKDDY	480
DB	495	DNQNLPSDNRIYDYSEYEDVTNDYQAQGNNDVNFQGNIDSPYIIKISKYDPNKDDY	554

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:49:18 ; Search time 68.491 Seconds  
(without alignments)  
2734.476 Million cell updates/sec

Title: US-10-806-288-13  
Perfect score: 3013  
Sequence: 1 SSDEKNDVINNQSIINTDD.....PDGTSKSVRTDEDKYQFDG 582

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pap:  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pap:  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pap:  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pap:  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pap:  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pap:  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pap:  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pap:  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pap:  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pap:  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pap:  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pap:  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pap:  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pap:  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pap:  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pap:  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pap:  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3013	100.0	1092	15	US-10-378-674-8
2	2810.5	93.3	892	12	US-10-282-122A-70481
3	2810.5	93.3	892	16	US-10-661-809-21
4	2810.5	93.3	930	12	US-10-615-383-10
5	2810.5	93.3	930	16	US-10-689-082-10
6	2810.5	93.3	930	16	US-10-690-184-10
7	2670	88.6	549	15	US-10-378-674-9
8	2510.5	83.3	560	15	US-10-378-674-2
9	1672	55.5	343	15	US-10-378-674-4
10	1562	51.8	316	15	US-10-378-674-6
11	1137.5	37.8	1141	12	US-10-282-122A-70251
12	1116.5	37.1	1470	12	US-10-282-122A-70444
13	576	19.1	1349	9	US-09-815-242-5898
14	576	19.1	1349	9	US-09-815-242-13137
15	559	18.6	1385	12	US-10-282-122A-44324

16 499.5 16.6 932 9 US-09-815-242-5578 Sequence 5578, Ap  
17 499.5 16.6 932 9 US-09-815-242-12438 Sequence 12438, A  
18 499 16.6 1742 12 US-10-615-383-4 Sequence 4, Appli  
19 499 16.6 1742 16 US-10-689-082-4 Sequence 4, Appli  
20 499 16.6 1742 16 US-10-690-184-4 Sequence 4, Appli  
21 493 16.4 1633 12 US-10-282-122A-70437 Sequence 70437, A  
22 489 16.2 953 12 US-10-282-122A-44457 Sequence 44457, A  
23 486.5 16.1 1920 12 US-10-282-122A-71413 Sequence 71413, A  
24 455 15.1 841 9 US-09-815-242-5779 Sequence 5779, Ap  
25 455 15.1 841 9 US-09-815-242-12751 Sequence 12751, A  
26 454 15.1 936 8 US-08-781-986A-5249 Sequence 5249, Ap  
27 454 15.1 936 12 US-10-329-624-5249 Sequence 5249, Ap  
28 452 15.0 935 12 US-10-282-122A-44326 Sequence 44326, A  
29 449 14.9 1021 9 US-09-815-242-5471 Sequence 5471, Ap  
30 449 14.9 1021 9 US-09-815-242-12544 Sequence 12544, A  
31 443.5 14.7 520 14 US-10-056-052-2 Sequence 2, Appli  
32 397.5 13.2 877 12 US-10-282-122A-70428 Sequence 70428, A  
33 377 12.5 496 14 US-10-311-879-29 Sequence 29, Appli  
34 376 12.5 767 9 US-09-815-242-5899 Sequence 5899, Ap  
35 376 12.5 767 9 US-09-815-242-13140 Sequence 13140, A  
36 349 11.6 331 14 US-10-056-052-4 Sequence 4, Appli  
37 349 11.6 345 9 US-09-813-820-7 Sequence 7, Appli  
38 344.5 11.4 961 12 US-10-282-122A-43778 Sequence 43778, A  
39 326 10.8 1038 12 US-10-282-122A-43827 Sequence 43827, A  
40 319.5 10.6 978 9 US-09-815-242-5456 Sequence 5456, Ap  
41 319.5 10.6 1001 9 US-09-815-242-12686 Sequence 12686, A  
42 308.5 10.2 1018 9 US-09-815-242-5797 Sequence 5797, Ap  
43 308.5 10.2 1018 9 US-09-815-242-12838 Sequence 12838, A  
44 300.5 10.0 1027 8 US-08-781-986A-5254 Sequence 5254, Ap  
45 300.5 10.0 1027 12 US-10-329-624-5254 Sequence 5254, Ap

## ALIGNMENTS

## RESULT 1

US-10-378-674-8  
; Sequence 8, Application US/10378674  
; Publication No. US20040006209A1  
; GENERAL INFORMATION:  
; APPLICANT: PARTI, Joseph M.  
; TITLE OF INVENTION: MONOCLONAL AND POLYCLONAL ANTIBODIES RECOGNIZING COAGULASE-NEGAT  
; FILE OF INVENTION: STAPHYLOCOCCAL PROTEINS  
; FILE REFERENCE: P07556US01/BAS  
; CURRENT APPLICATION NUMBER: US/10/378,674  
; CURRENT FILING DATE: 2003-03-05  
; PRIOR APPLICATION NUMBER: 60/361,324  
; PRIOR FILING DATE: 2002-03-05  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 8  
; LENGTH: 1092  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-10-378-674-8

Query Match 100.0%; Score 3013; DB 15; Length 1092;  
Best Local Similarity 100.0%; Pred. No. 1.3e-176;  
Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSDEKNDVINNQSIINTDDNNQIIKKETNNYDGIKESDRTSTTNVDNEATFLOK 60  
DB 75 SSDEKNDVINNQSIINTDDNNQIIKKETNNYDGIKESDRTSTTNVDNEATFLOK 134  
QY 61 TPQDNTHLTEREEVKSSSVESNSSIDTAQPSHTTINREESVQSDNVEDSHVDFANS 120  
DB 135 TPQDNTHLTEREEVKSSSVESNSSIDTAQPSHTTINREESVQSDNVEDSHVDFANS 194  
QY 121 KIKESNTSGKEENTIEQPNKVEDSTTSQPSGYTNIDEKISNQDELLNLPINEYENKAR 180  
DB 195 KIKESNTSGKEENTIEQPNKVEDSTTSQPSGYTNIDEKISNQDELLNLPINEYENKAR 254  
QY 181 PLSTTSQAQPSIKRVTNNQLAAEQGSNVNHLIKVTOQSTEGYDDSEGVKAHAENLIYD 240

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:44:57 ; Search time 17.0358 Seconds  
(without alignments)  
1763.712 Million cell updates/sec

Title: US-10-806-288-13  
Perfect score: 3013  
Sequence: 1 SSDEKNDVINNQSIINTDD.....PDGTSKSVRTDCKGYQFDG 582

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgm2\_6/ptodata/2/iaa/5A.COMB.pdp:\*  
2: /cgm2\_6/ptodata/2/iaa/5B.COMB.pdp:\*  
3: /cgm2\_6/ptodata/2/iaa/6A.COMB.pdp:\*  
4: /cgm2\_6/ptodata/2/iaa/6B.COMB.pdp:\*  
5: /cgm2\_6/ptodata/2/iaa/PTUS.COMB.pdp:\*  
6: /cgm2\_6/ptodata/2/iaa/backfiles1.pdp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2810.5	93.3	930	4	US-09-134-001C-5314
2	2810.5	93.3	930	4	US-09-386-962C-10
3	1144.5	38.0	1166	4	US-09-200-650B-7
4	576	19.1	1315	4	US-09-200-650B-5
5	499	16.6	1742	4	US-09-386-962C-4
6	455	15.1	930	4	US-09-200-650B-3
7	454	15.1	936	4	US-08-956-171E-5249
8	452	15.0	933	3	US-08-293-728-2
9	452	15.0	933	3	US-09-421-868-2
10	396	13.1	918	4	US-09-200-650B-1
11	349	11.6	345	3	US-08-856-253-7
12	300.5	10.0	1027	4	US-08-956-171E-5254
13	259.5	8.6	1112	2	US-08-714-402-2
14	259.5	8.6	1161	4	US-09-327-536-2
15	220.5	7.3	3696	4	US-09-134-001C-5080
16	209	6.9	630	3	US-08-973-462-9
17	208.5	6.9	251	4	US-08-956-171E-5252
18	203.5	6.8	2504	4	US-09-328-352-5921
19	202.5	6.7	368	4	US-09-134-000C-4597
20	202	6.7	676	4	US-09-134-001C-4318
21	201.5	6.7	1177	4	US-09-134-001C-5106
22	197	6.5	2314	4	US-09-268-347-49
23	195	6.5	1335	4	US-09-134-001C-3716
24	195	6.5	10182	4	US-09-134-001C-3159
25	192.5	6.4	1306	4	US-09-134-000C-6670
26	192.5	6.4	1786	3	US-08-973-462-8
27	190.5	6.3	1833	4	US-08-621-944A-4

28	190.5	6.3	1833	4	US-08-945-567D-4	Sequence 4, Appli
29	190.5	6.3	1992	4	US-08-621-944A-3	Sequence 3, Appli
30	190.5	6.3	1992	4	US-08-945-567D-3	Sequence 3, Appli
31	189.5	6.3	1664	1	US-09-599-652-2	Sequence 2, Appli
32	189.5	6.3	1664	2	US-08-642-846-2	Sequence 2, Appli
33	189.5	6.3	1664	4	US-09-264-604-2	Sequence 2, Appli
34	188.5	6.3	1435	2	US-08-568-459A-4	Sequence 4, Appli
35	188.5	6.3	1435	2	US-08-487-826B-4	Sequence 4, Appli
36	188.5	6.3	1435	4	US-09-210-288-4	Sequence 4, Appli
37	187	6.2	1416	4	US-09-071-035-404	Sequence 404, App
38	187	6.2	1448	4	US-09-071-035-402	Sequence 402, App
39	186.5	6.2	3088	4	US-09-562-702A-8	Sequence 8, Appli
40	186.5	6.2	3089	4	US-09-562-702A-4	Sequence 4, Appli
41	186.5	6.2	3110	4	US-09-562-702A-2	Sequence 2, Appli
42	186.5	6.2	3110	4	US-09-562-702A-6	Sequence 6, Appli
43	186.5	6.2	3110	4	US-09-561-709B-7	Sequence 7, Appli
44	186.5	6.2	3111	2	US-08-460-309-4	Sequence 4, Appli
45	186.5	6.2	3111	2	US-08-125-077-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1  
US-09-134-001C-5314  
; Sequence 5314, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GPC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; PRIOR FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 5314  
; LENGTH: 930  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-5314

Query Match 93.3%; Score 2810.5; DB 4; Length 930;  
Best Local Similarity 93.6%; Pred. No. 2e-158;  
Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;

Qy	1	SSDEKNDVINNQSIINTDDNNQIIKKEETNNYDGIKRSSEDRTESTTNVDENEATFLQK 60	
Db	73	SSNEEKNDVINNQSIINTDDNQ-IKKEETNSDAIENRSKDIQTSTTNVDENEATFLQK 131	
Qy	61	TPQNTLHTEBEVKSSSVSSSSSIDTAQPSHTTNRESVOTSDNVDSHVSDFPANS 120	
Db	132	TPQNTQLKEEVKPSVSSSSSMVTAQPSHTTNSEASIQTSDEKNSRVSDPANS 191	
Qy	121	KIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNIDEKISNQDELNLPINEYENKAR 180	
Db	192	KIIESNTSKEENTIEQPNKVRSDTSQPSGYKNIDEKISNQDELNLPINEYENKVR 251	
Qy	181	PLSTTSQPSIKRVTVNQLAAQGSNNVHLIKVTDQISITEGYDSDSEGVIKHAHAENLIYD 240	
Db	252	PLSTTSQPSIKRVTVNQLAAQGSNNVHLIKVTDQISITEGYDSDSEGVIKHAHAENLIYD 311	
Qy	241	VTFEVDKVKSGDVTMTVDIKNTVPSDLTDSFTTPKIKNSGEIATCTYDNKNQIYTT 300	
Db	312	VTFEVDKVKSGDVTMTVDIKNTVPSDLTDSFTTPKIKNSGEIATCTYDNKNQIYTT 371	
Qy	301	FTDVKYENIKAHKLKLTSTYIDSKVPNNNTKLDVEYKNTALSSVKNKTTIVYQRPENRT 360	
Db	372	FTDVKYENIKAHKLKLTSTYIDSKVPNNNTKLDVEYKNTALSSVKNKTTIVYQRPENRT 431	

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:33:49 ; Search time 65.362 Seconds  
(without alignments)  
2515.875 Million cell, updates/sec

Title: US-10-806-288-13  
Perfect score: 3013  
Sequence: 1 SSDEKNDVINNQSIINTDD.....PDGTSKSVRTDEGKQFQDG 582

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the chance being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3013	100.0	1092	2 AAW41602	Aaw41602 Staphyloc
2	3013	100.0	1092	7 ABM79019	Abm79019 Staphyloc
3	2810.5	93.3	892	6 ABU42557	Abu42557 Protein e
4	2810.5	93.3	930	5 ABP40469	Abp40469 Staphyloc
5	2810.5	93.3	991	3 AAY83171	Aay83171 Cell wall
6	2810.5	93.3	991	3 AAY70120	Aay70120 Staph. ep
7	2670	88.6	549	7 ABM79020	Abm79020 Staphyloc
8	2510.5	83.3	560	7 ABM79015	Abm79015 Staphyloc
9	1672	55.5	343	7 ABM79016	Abm79016 Staphyloc
10	1562	51.8	316	7 ABM79017	Abm79017 Staphyloc
11	1144.5	38.0	1166	2 AAY08643	Aay08643 S. aureus
12	1142.5	37.9	1166	6 ABJ18982	Abj18982 Pathogen
13	1137.5	37.8	1141	6 ABU42327	Abu42327 Protein e
14	1116.5	37.1	670	6 ABU42520	Abu42520 Protein e
15	857	28.4	278	4 AAG82803	Aag82803 S. epider
16	576	19.1	1315	2 AAY08642	Aay08642 S. aureus
17	576	19.1	1315	6 ABJ18969	Abj18969 Pathogen
18	576	19.1	1349	4 AAY37544	Aay37544 Staphyloc
19	576	19.1	1349	4 AAY34402	Aay34402 Staphyloc
20	576	19.1	1349	6 ABM72436	Abm72436 Staphyloc
21	559	18.6	1385	6 ABU16400	Abu16400 Protein e
22	499.5	16.6	932	4 AAY36845	Aay36845 Staphyloc
23	499.5	16.6	932	4 AAY34082	Aay34082 Staphyloc
24	499	16.6	1802	3 AAY83170	Aay83170 Cell wall
25	499	16.6	1802	3 AAY70119	Aay70119 Staph. ep

## ALIGNMENTS

RESULT 1  
AAW41602

ID AAW41602 standard; protein; 1092 AA.

AC AAW41602;

XX XX

DT 17-OCT-2003 (revised)

DT 22-JUN-1998 (first entry)

XX Staphylococcus epidermidis fibrinogen binding protein FIG.

DE Fibrinogen binding protein; FIG; aggregation; infection;

KW coagulase-negative Staphylococcus; therapy; diagnosis; immunisation;

KW immunogen; vaccine.

OS Staphylococcus epidermidis; strain HB.

XX FH Key Location/Qualifiers

FT Peptide 1..51

FT Protein /label= Sig\_peptide

FT Region /label= Mat\_protein

FT Region 52..824

FT Region /note= "non-repetitive region, harbours fibrinogen binding activity"

FT Region 825..1040

FT Region /note= "Asp-Ser dipeptide repeat region"

FT Region 1053..1057

FT Region /note= "cell wall anchoring motif"

XX WO9748727-A1.

PN 24-DEC-1997.

XX 18-JUN-1997; 97WO-SE001091.

XX 20-JUN-1996; 96SE-00002496.

XX (GUSL)/ GUSL B.

XX (NILS)/ NILSSON M.

XX (FRYK)/ FRYKBERG L.

XX (FLOC)/ FLOCK J.

XX (LIND)/ LINDBERG M.

XX Guss B, Nilsson M, Frykberg L, Flock J, Lindberg M;

XX WPI; 1998-063079/06.

XX N-FSDB; AAV04279.

Aag82343 S. epider  
Abu42513 Protein e  
Abu16533 Protein e  
Abu43489 Protein e  
Aau37158 Staphyloc  
Aau34283 Staphyloc  
Aay08641 S. aureus  
Abj18947 Pathogen  
Abj18940 Pathogen  
Abm72437 Staphyloc  
Abm72221 Staphyloc  
Aaw9801 Staphyloc  
Aay58435 Staphyloc  
Aab69508 Staphyloc  
Abu16402 Protein e  
Aau33975 Staphyloc  
Aau36951 Staphyloc  
Aae29262 Staphyloc  
Ada89555 Staphyloc  
Ada89539 Staphyloc



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:33:49 ; Search time 65.362 Seconds  
(without alignments)  
2515.875 Million cell updates/sec

Title: US-10-806-288-13  
Perfect score: 3013  
Sequence: 1 SSDEEKNDVINNQSNITDD.....PDGTSKSVRTDEDKYQFDG 582

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3013	100.0	1092	2 AAW41602	Aaw41602 Staphyloc
2	3013	100.0	1092	7 ABM79019	Abm79019 Staphyloc
3	2810.5	93.3	892	6 ABU42557	Abu42557 Protein e
4	2810.5	93.3	930	5 ABP40469	Abp40469 Staphyloc
5	2810.5	93.3	991	3 AAY83171	Aay83171 Cell wall
6	2810.5	93.3	991	3 AAY70120	Aay70120 Staph. ep
7	2670	88.6	549	7 ABM79020	Abm79020 Staphyloc
8	2510.5	83.3	560	7 ABM79015	Abm79015 Staphyloc
9	1672	55.5	343	7 ABM79016	Abm79016 Staphyloc
10	1562	51.8	316	7 ABM79017	Abm79017 Staphyloc
11	1144.5	38.0	1166	2 AAY08643	Aay08643 S. aureus
12	1142.5	37.9	1166	6 ABJ18982	Abj18982 Pathogen
13	1137.5	37.8	1141	6 ABU42327	Abu42327 Protein e
14	1116.5	37.1	670	6 ABU42520	Abu42520 Protein e
15	857	28.4	278	4 AAG82803	Aag82803 S. epider
16	576	19.1	1315	2 AAY08642	Aay08642 S. aureus
17	576	19.1	1315	6 ABJ18969	Abj18969 Pathogen
18	576	19.1	1349	4 AAU37544	Aau37544 Staphyloc
19	576	19.1	1349	4 AAU34402	Aau34402 Staphyloc
20	576	19.1	1349	6 ABM72436	Abm72436 Staphyloc
21	559	18.6	1385	6 ABU16400	Abu16400 Protein e
22	499.5	16.6	932	4 AAU36845	Aau36845 Staphyloc
23	499.5	16.6	932	4 AAU34082	Aau34082 Staphyloc
24	499	16.6	1802	3 AAY83170	Aay83170 Cell wall
25	499	16.6	1802	3 AAY70119	Aay70119 Staph. ep

26	497	16.5	1155	4 AAG82343	Aag82343 S. epider
27	493	16.4	1633	6 ABU42513	Abu42513 Protein e
28	489	16.2	953	6 ABU16533	Abu16533 Protein e
29	486.5	16.1	1920	6 ABU43489	Abu43489 Protein e
30	455	15.1	841	4 AAU37158	Aau37158 Staphyloc
31	455	15.1	841	4 AAU34283	Aau34283 Staphyloc
32	455	15.1	930	2 AAY08641	Aay08641 S. aureus
33	455	15.1	933	6 ABJ18947	Abj18947 Pathogen
34	455	15.1	947	6 ABU18940	Abu18940 Pathogen
35	455	15.1	995	6 ABM72437	Abm72437 Staphyloc
36	454	15.1	927	6 ABM72221	Abm72221 Staphyloc
37	454	15.1	936	2 AAW89801	Aaw89801 Staphyloc
38	452	15.0	933	3 AAY58435	Aay58435 Staphyloc
39	452	15.0	933	4 AAB89508	Aab89508 Staphyloc
40	452	15.0	935	6 ABU16402	Abu16402 Protein e
41	449	14.9	1021	4 AAU33975	Aau33975 Staphyloc
42	449	14.9	1021	4 AAU36951	Aau36951 Staphyloc
43	443.5	14.7	520	5 AAE29262	Aae29262 Staphyloc
44	397.5	13.2	567	6 ADA89555	Ada89555 Staphyloc
45	397.5	13.2	877	6 ADA89539	Ada89539 Staphyloc

## ALIGNMENTS

RESULT 1  
AAW41602  
ID AAW41602 standard; protein; 1092 AA.  
XX  
AC AAW41602;  
XX  
DT 17-OCT-2003 (revised)  
DT 22-JUN-1998 (first entry)  
XX  
DE Staphylococcus epidermidis fibrinogen binding protein FIG.  
XX  
KW Fibrinogen binding protein; FIG; aggregation; infection;  
KW coagulase-negative staphylococcus; therapy; diagnosis; immunisation;  
KW immunogen; vaccine.  
XX  
OS Staphylococcus epidermidis; strain HB.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..51  
FT Protein /label= Sig\_peptide 52..1092  
FT Protein /label= Mat\_protein 52..824  
FT Region /note= "non-repetitive region, harbours fibrinogen binding activity" 825..1040  
FT Region /note= "Asp-Ser dipeptide repeat region" 1053..1057  
FT Region /note= "cell wall anchoring motif"  
XX WO9748727-A1.  
PN  
XX 24-DEC-1997.  
XX 18-JUN-1997; 97WO-SE001091.  
XX 20-JUN-1996; 96SE-00002496.  
XX (GUSS/) GUSS B.  
XX (NILS/) NILSSON M.  
XX (FRYK/) FRYKBERG L.  
XX (FLOC/) FLOCK J.  
XX (LIND/) LINDBERG M.  
XX Guss B, Nilsson M, Frykberg L, Flock J, Lindberg M;  
XX WPI; 1998-063079/06.  
XX N-PSDB; AAV04279.

XX Fibrinogen-binding protein from coagulase-negative Staphylococcus - used  
 PT for prevention, treatment and diagnosis of Staphylococcus infection.  
 XX  
 PS Example 3; Fig 6; 45pp; English.  
 XX  
 CC The protein comprises the fibrinogen binding protein (FIG) of coagulase-  
 CC negative Staphylococcus epidermidis HB. Its amino acid sequence was  
 CC deduced from the isolated fig gene (see AAV04279). The closest known  
 CC analogue of FIG is the clumping factor of *S. aureus* which also binds  
 CC fibrinogen and promotes bacterial aggregation in serum. Recombinant FIG  
 CC polypeptides can be expressed in host cells. They are used as immunogens,  
 CC particularly in vaccines (which may be expressed in vivo) to protect  
 CC humans and animals against coagulase-negative Staphylococcus infection.  
 CC Antibodies raised against FIG can be used for passive immunisation. They  
 CC block the adherence of bacteria) and for diagnosis. (Updated on 17-OCT-  
 CC 2003 to standardise OS field)  
 XX  
 SQ Sequence 1092 AA;

Query Match 100.0%; Score 3013; DB 2; Length 1092;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-156;  
 Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SSDEKDVNNQSDINTDDNNQIIKKEETNNYDGIKESDRTSTNNVDENEAFLQK 60  
 DB 75 SSDEKDVNNQSDINTDDNNQIIKKEETNNYDGIKESDRTSTNNVDENEAFLQK 134  
 QY 61 TPQDNTHLTEEEVKSSSVESNSSIDTAQOPSHTTINREESVQTSNVEDSHVSDFANS 120  
 DB 135 TPQDNTHLTEEEVKSSSVESNSSIDTAQOPSHTTINREESVQTSNVEDSHVSDFANS 194  
 QY 121 KIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQDELNLNLPINEYENKAR 180  
 DB 195 KIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQDELNLNLPINEYENKAR 254  
 QY 181 PLSTTSAQPSIKRVTNQLAAEQGSNNVHLIKVTDQSITEGYDDSEGVKAHAENLIYD 240  
 DB 255 PLSTTSAQPSIKRVTNQLAAEQGSNNVHLIKVTDQSITEGYDDSEGVKAHAENLIYD 314  
 QY 241 VTFEVDDKVKSGDWTVDIDKNTVPDLTDSFTPIPKIKNSGEIATGYDNKNKQIYTT 300  
 DB 315 VTFEVDDKVKSGDWTVDIDKNTVPDLTDSFTPIPKIKNSGEIATGYDNKNKQIYTT 374  
 QY 301 FTDYVDKYENIKAHKLKLTSTYIDKSKVPNNNTKLDVEYKTLSSVNTITVEYQRPENRT 360  
 DB 375 FTDYVDKYENIKAHKLKLTSTYIDKSKVPNNNTKLDVEYKTLSSVNTITVEYQRPENRT 434  
 QY 361 ANLQSMFTNIDTKNHTVEQTIYINPLYSKATETNNVINSNGDEGSTIIDSTIIKVKYKVG 420  
 DB 435 ANLQSMFTNIDTKNHTVEQTIYINPLYSKATETNNVINSNGDEGSTIIDSTIIKVKYKVG 494  
 QY 421 DNQNLPSNRNIYDYSEYEDVTNDYDQAQLGNNDVNFNFGNIDSPYIIKVISKYPNKKDY 480  
 DB 495 DNQNLPSNRNIYDYSEYEDVTNDYDQAQLGNNDVNFNFGNIDSPYIIKVISKYPNKKDY 554  
 QY 481 TTIQQTVMQTINNEYTGFEPTASNDNTIAPSTSSGQGGDLPEPKTYKIGDYVWEDVK 540  
 DB 555 TTIQQTVMQTINNEYTGFEPTASNDNTIAPSTSSGQGGDLPEPKTYKIGDYVWEDVK 614  
 QY 541 DGIQNTNDNEKPLSNVLTLVTPDGTSKSVRTDDEGKYQFDG 582  
 DB 615 DGIQNTNDNEKPLSNVLTLVTPDGTSKSVRTDDEGKYQFDG 656

RESULT 2

ABM79019

ID ABM79019 standard; protein; 1092 AA.

XX

AC ABM79019;

XX

DT 15-JAN-2004 (first entry)

XX

DE Staphylococcus epidermidis polypeptide.  
 XX Infection; antibacterial; vaccine.  
 XX

OS Staphylococcus epidermidis.  
 XX

PN WO2003076470-A1.  
 XX

PD 18-SEP-2003.  
 XX

PF 05-MAR-2003; 2003WO-US006415.  
 XX

PR 05-MAR-2002; 2002US-0361324P.  
 XX

XX (INHIBITEX INC.  
 PA (TEXA ) UNIV TEXAS A & M SYSTEM.

XX

PI Patti JM, Hutchins JT, Hall A, Domanski P, Patel P, Hook M;  
 PT Robbins J, Vernachio J, Bowden MG;

XX WPI; 2003-722324/68.

XX

PT New antibody recognizing a Staphylococcus epidermidis protein comprising  
 PT SdrG NIN23, SdrG N23 or SdrGTR2 useful for preparing a composition for  
 PT treating or preventing a coagulase-negative Staphylococcus infection.  
 XX

PS Claim 16; Page 36-37; 78pp; English.

XX

CC The present sequence comprises the protein sequence of a polypeptide of a  
 CC coagulase-negative Staphylococcus epidermidis. A claimed monoclonal  
 CC antibody recognises this protein and is used in a claimed method of  
 CC treating or preventing a coagulase-negative staphylococcus infection in a  
 CC human or animal, e.g. a nosocomial coagulase-negative staphylococcus  
 CC infection in low birth weight infants  
 XX

SQ Sequence 1092 AA;

Query Match 100.0%; Score 3013; DB 7; Length 1092;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-156;  
 Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSDEKDVNNQSDINTDDNNQIIKKEETNNYDGIKESDRTSTNNVDENEAFLQK 60

DB 75 SSDEKDVNNQSDINTDDNNQIIKKEETNNYDGIKESDRTSTNNVDENEAFLQK 134

QY 61 TPQDNTHLTEEEVKSSSVESNSSIDTAQOPSHTTINREESVQTSNVEDSHVSDFANS 120

DB 135 TPQDNTHLTEEEVKSSSVESNSSIDTAQOPSHTTINREESVQTSNVEDSHVSDFANS 194

QY 121 KIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQDELNLNLPINEYENKAR 180

DB 195 KIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQDELNLNLPINEYENKAR 254

QY 181 PLSTTSAQPSIKRVTNQLAAEQGSNNVHLIKVTDQSITEGYDDSEGVKAHAENLIYD 240

DB 255 PLSTTSAQPSIKRVTNQLAAEQGSNNVHLIKVTDQSITEGYDDSEGVKAHAENLIYD 314

QY 241 VTFEVDDKVKSGDWTVDIDKNTVPDLTDSFTPIPKIKNSGEIATGYDNKNKQIYTT 300

DB 315 VTFEVDDKVKSGDWTVDIDKNTVPDLTDSFTPIPKIKNSGEIATGYDNKNKQIYTT 374

QY 301 FTDYVDKYENIKAHKLKLTSTYIDKSKVPNNNTKLDVEYKTLSSVNTITVEYQRPENRT 360

DB 375 FTDYVDKYENIKAHKLKLTSTYIDKSKVPNNNTKLDVEYKTLSSVNTITVEYQRPENRT 434

QY 361 ANLQSMFTNIDTKNHTVEQTIYINPLYSKATETNNVINSNGDEGSTIIDSTIIKVKYKVG 420

DB 435 ANLQSMFTNIDTKNHTVEQTIYINPLYSKATETNNVINSNGDEGSTIIDSTIIKVKYKVG 494

QY 421 DNQNLPSNRNIYDYSEYEDVTNDYDQAQLGNNDVNFNFGNIDSPYIIKVISKYPNKKDY 480

DB 495 DNQNLPSNRNIYDYSEYEDVTNDYDQAQLGNNDVNFNFGNIDSPYIIKVISKYPNKKDY 554

Qy 481 TTIOQTVMQTTINEYTGFEPTASVDNTIAPSTSSGQGGDLPPKTYKIGDYVWEDVK 540  
Db 555 TTIOQTVMQTTINEYTGFEPTASVDNTIAPSTSSGQGGDLPPKTYKIGDYVWEDVK 614  
Qy 541 DGIQNTNDNEKPLSNVLTLTYPDGTSKSVRTDDEGKYQFDG 582  
Db 615 DGIQNTNDNEKPLSNVLTLTYPDGTSKSVRTDDEGKYQFDG 656

## RESULT 3

ABU42557  
ID ABU42557 standard; protein; 892 AA.

XX AC ABU42557;

XX XX 19-JUN-2003 (first entry)

XX XX Protein encoded by Prokaryotic essential gene #28084.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Staphylococcus epidermidis.

XX PN WO200277183-A2.

XX PD 03-OCT-2002.

XX XX 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
XX PT Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX DR WPI; 2003-029926/02.

XX DR N-PSDB; ACA46427.

XX PT New antisense nucleic acids, useful for identifying proteins or screening  
XX PT for homologous nucleic acids required for cellular proliferation to  
XX PT isolate candidate molecules for rational drug discovery programs.

XX PS Claim 25; SEQ ID NO 70481; 1766pp; English.

XX CC The invention relates to an isolated nucleic acid comprising any one of  
XX CC the 6213 antisense sequences given in the specification where expression  
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid  
XX CC encoding a polypeptide whose expression is inhibited by the antisense  
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
XX CC polypeptide or its fragment whose expression is inhibited by the  
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding  
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
XX CC proliferation or the activity of a gene in an operon required for  
XX CC proliferation; (7) identifying a compound that influences the activity of  
XX CC the gene product or that has an activity against a biological pathway  
XX CC required for proliferation, or that inhibits cellular proliferation; (8)  
XX CC identifying a gene required for cellular proliferation or the biological  
XX CC pathway in which a proliferation-required gene or its gene product lies  
XX CC or a gene on which the test compound that inhibits proliferation of an  
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
XX CC compound's activity; (11) a culture comprising strains in which the gene  
XX CC product is overexpressed or underexpressed; (12) determining the extent  
XX CC to which each of the strains is present in a culture or collection of  
XX CC strains; or (13) identifying the target of a compound that inhibits the  
XX CC proliferation of an organism. The antisense nucleic acids are useful for  
XX CC identifying proteins or screening for homologous nucleic acids required  
XX CC for cellular proliferation to isolate candidate molecules for rational

CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 892 AA;

Query Match 93.3%; Score 2810.5; DB 6; Length 892;  
Best Local Similarity 93.6%; Pred. No. 23e-145;  
Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;

Qy 1 SSDEKNDVINNQSIINTDDNQIIKKEETNNYDGIKRSERDSTSTNNVDENAEATFLQK 60  
Db 35 SSNEEKNDVINNQSIINTDDNQ-IKKEETNSDAIENRSKIDITQSTNNVDENAEATFLQK 93

Qy 61 TPQDNTLHTEBEVKESSESSSSSDTAQPSHTTINREESVQTSNVEDSHVDFANS 120  
Db 94 TPQDNTQLKEEVKPESSSESSSSMDTAQPSHTTINSEASIQTSNDEENSRVDFANS 153

Qy 121 KIKESNTESGKEENTIEQPNKVEDSTTSQPSGVNIDEKISNODELNLPINEYENKAR 180  
Db 154 KIKESNTESGKEENTIEQPNKVEDSTTSQPSGVNIDEKISNODELNLPINEYENKAR 213

Qy 181 PLSTTSAQPSIKRVTVNQLAAEQGSNNVHLIKVTDQSITEGYDDSEGVKAHAENLIYD 240  
Db 214 PLSTTSAQPSIKRVTVNQLAAEQGSNNVHLIKVTDQSITEGYDDSGIKAHAENLIYD 273

Qy 241 VTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFPIPKIKNSGELIATGTDNKNKQIYTT 300  
Db 274 VTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFPIPKIKNSGELIATGTDNKNKQIYTT 333

Qy 301 FTDVVDKYENIKAKHLKLTYSIDKSKVPNNNTKLDVEYKLTALSSVKNKTIIVYEQRPENRT 360  
Db 334 FTDVVDKYENIKAKHLKLTYSIDKSKVPNNNTKLDVEYKLTALSSVKNKTIIVYEQRPENRT 393

Qy 361 ANLQSMFTNIDTKNHTVQETIYINPLRYSAKETNNVINSNGDEGSTIIDSTIIKVKVG 420  
Db 394 ANLQSMFTNIDTKNHTVQETIYINPLRYSAKETNNVINSNGDEGSTIIDSTIIKVKVG 453

Qy 421 DNQNLPSNRIRIYDSEYEDVTNDDYAQLGNNDVINFGNIDSPYIIKVISKYDPNKDDY 480  
Db 454 DNQNLPSNRIRIYDSEYEDVTNDDYAQLGNNDVINFGNIDSPYIIKVISKYDPNKDDY 513

Qy 481 TTIOQTVMQTTINEYTGFEPTASVDNTIAPSTSSGQGGDLPPKTYKIGDYVWEDVK 540  
Db 514 TTIOQTVMQTTINEYTGFEPTASVDNTIAPSTSSGQGGDLPPKTYKIGDYVWEDVK 573

Qy 541 DGIQNTNDNEKPLSNVLTLTYPDGTSKSVRTDDEGKYQFDG 582  
Db 574 DGIQNTNDNEKPLSNVLTLTYPDGTSKSVRTDDEGKYQFDG 615

RESULT 4  
ABP40469

ID ABP40469 standard; protein; 930 AA.

XX AC ABP40469;

XX XX 24-JUL-2002 (first entry)

XX DT Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:5314.

XX DE Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;

XX KW antibacterial; gene therapy.

XX OS Staphylococcus epidermidis.

XX XX US6380370-B1.

XX XX 30-APR-2002.

XX PD

Db		612	DGIQNTNDNEKPLSNVLVTLPDGTSGSVRTDEBGKYQFDG	53
	RESULT 5			
	AAy83171			
ID	AAy83171	standard; protein;	991 AA.	
XX	XX			
AC	AAy83171,			
XX	XX			
DT	24-JUL-2000	(first entry)		
XX	XX			
DE	Cell wall protein SdrG.			
XX	XX			
KW	SdrF, SdrG; SdrH; coagulase negative; staphylococcus;			
KW	osteomyelitis; endocarditis; immune response; vaccine;			
KW	intravenous catheter; heart valve; cardiac.			
XX	XX			
OS	Staphylococcus sp.			
XX	XX			
FH	Key	Location/Qualifiers		
FT	Misc-difference 14	/note= "Position encoded by TAG stop codon"		
FT	Misc-difference 33	/note= "Position encoded by TGA stop codon"		
FT	Misc-difference 94	/note= "Position encoded by TAA stop codon"		
FT	Misc-difference 980	/note= "Position encoded by TAG stop codon"		
FT	Misc-difference 989	/note= "Position encoded by TAA stop codon"		
FT	FT			
FT	FT			
FT	FT			

Query Match	93.3%	Score 2810.5;	DB 5;	Length 930;
Best Local Similarity	93.6%;	Pred. No. 2.4e-145;		
Matches 545;	Conservative 15;	Mismatches 21;	Indels 11;	Gaps 1;

[illegible]

QY 1 SSDEEKNDVINNNQSINTDDNNQIIKKEETNNYDGIKRSRDETRSTNNVDNEATFLQK 60

Db 106 SSNEKNDVINNSQSIINTDDNQ-IKKEETNSDAIENRSKDIQTQSTTNVDENEATFLQK 164  
 Qy 61 TPQDNTHLTBEVKESSVSSNSIDTAQPSHTTINREESVQTSNDVDSHVSDFANS 120  
 Db 165 TPQDNTHLTBEVKESSVSSNSIDTAQPSHTTINSEASIQTSNEENSRSVDFANS 224  
 Qy 121 KIKESNTSGKEENTIEQPNKVKEDSTSPQSGVTNIDKISNODELNLNPINEYENKAR 180  
 Db 225 KIKESNTSGKEENTIEQPNKVKEDSTSPQSGVTNIDKISNODELNLNPINEYENKAR 284  
 Qy 181 PLSTTSAPSKRVTVNQLAAEQGSNNVHLIKVTDQSIETEGYDDSEGVIKAHAENLIYD 240  
 Db 285 PLSTTSAPSKRVTVNQLAAEQGSNNVHLIKVTDQSIETEGYDDSEGVIKAHAENLIYD 344  
 Qy 241 VTPEVDDKVKSGDGTMTVDIDKNTVPSDLTDSFTPIKIDNSGEIATGTVDNKNKQIYTT 300  
 Db 345 VTPEVDDKVKSGDGTMTVDIDKNTVPSDLTDSFAIPKIDNSGEIATGTVDNKNKQIYTT 404  
 Qy 301 FTDYVDKVENIKAHKLKTSYIDKSKVPNNNTKLDVEYKLTALSSVNTKITVEYQRPENRT 360  
 Db 405 FTDYVDKVENIKAHKLKTSYIDKSKVPNNNTKLDVEYKLTALSSVNTKITVEYQRPENRT 464  
 Qy 361 ANLOSMTNIDTKNHTVEQTIYINPLYSKATETNWNISGNGDEGTTIIDSTIIKVKVG 420  
 Db 465 ANLOSMTNIDTKNHTVEQTIYINPLYSKATETNWNISGNGDEGTTIIDSTIIKVKVG 524  
 Qy 421 DNQNLPSNRIDYSEYEDVTNDYDQAGNNNDVNIINFGNIDSPYIIKVISKYDPNKDDY 480  
 Db 525 DNQNLPSNRIDYSEYEDVTNDYDQAGNNNDVNIINFGNIDSPYIIKVISKYDPNKDDY 584  
 Qy 481 TTIQQTVMQTTINEYTGFEFTASVDNTIAESTSGGQGGDLPEKYIKIGDYWEDVDK 540  
 Db 585 TTIQQTVMQTTINEYTGFEFTASVDNTIAESTSGGQGGDLPEKYIKIGDYWEDVDK 644  
 Qy 541 DGIONTNDNEPLSNVLTLVYPDGTSKSVRTDGGKVPQDG 582  
 Db 645 DGIONTNDNEPLSNVLTLVYPDGTSKSVRTDGGKVPQDG 686

## RESULT 6

AA70120  
 ID AA70120 standard; protein; 991 AA.

XX AC AA70120;  
 XX DT

XX 06-JUN-2000 (first entry)

XX Staph. epidermidis serine-aspartate repeat region protein SdrG.

XX Multicomponent vaccine; immunostimulatory; antibacterial; MSCRAMM;  
 KW microbial surface components recognising adhesive matrix molecules;  
 KW collagen binding protein; CBP; CNA; fibrinogen binding protein;  
 KW Clumping factor A; ClfA; Clumping factor B; ClfB; FnBP;  
 KW fibronectin binding protein; Staphylococcus infection;  
 KW serine-aspartate repeat region protein; SDR protein; SdrG.

XX Staphylococcus epidermidis.

OS Key Location/Qualifiers  
 FH Misc-difference 14 /note= "Encoded by in-frame stop codon TAG"  
 FT Misc-difference 33 /note= "Encoded by in-frame stop codon TGA"  
 FT Misc-difference 964 /note= "Encoded by in-frame stop codon TAA"  
 FT Misc-difference 980 /note= "Encoded by in-frame stop codon TAG"  
 FT Misc-difference 989 /note= "Encoded by in-frame stop codon TAA"

XX WO200012131-A1.

XX 09-MAR-2000.

XX 31-AUG-1999; 99WO-US019727.  
 XX 31-AUG-1998; 98US-0098439P.  
 XX (INH-) INHIBITEX INC.  
 PA (TEXA ) UNIV TEXAS A & M SYSTEM.  
 PA (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.  
 XX Patti JM, Foster TJ, Hook M;  
 XX MPI; 2000-237781/20.  
 DR N-PSDB; AAZ51202.  
 XX Composition used for generating immune response or for inhibiting  
 PT microbial colonization in an animal comprises antibodies that bind  
 PT collagen binding protein, fibrinogen binding protein and, optionally,  
 PT fibronectin binding protein.  
 XX Claim 8; Fig 4; 115pp; English.  
 XX The patent discloses multicomponent vaccines containing selected  
 CC combinations of bacterial binding proteins termed MSCRAMM (microbial  
 CC surface components recognising adhesive matrix molecules) or their  
 CC antibodies. A vaccine composition is provided that includes collagen  
 CC binding protein or peptide, e.g. CNA, a fibrinogen binding protein  
 CC preferably clumping factor A (ClfA) or Clumping factor B (ClfB), and  
 CC optionally a fibronectin binding protein e.g. FnBP-A. The vaccines are  
 CC useful for imparting protection against a broad spectrum of  
 CC Staphylococcal strains and for inhibiting microbial colonisation,  
 CC especially of Staphylococcus aureus, in an animal. The combinations can  
 CC also be used to select donor blood pools for the preparation of purified  
 CC blood products for passive immunisation. The present sequence is a serine  
 CC -aspartate repeat region protein, SdrG from Staphylococcus epidermidis.  
 CC The Sdr protein is useful in vaccine preparation in combination with  
 CC specific bacterial binding proteins. These vaccines can be used to treat  
 CC a broad spectrum of bacterial infections, including those arising from  
 CC both coagulase-positive and coagulase-negative bacteria  
 XX SQ Sequence 991 AA;

Query Match 93.3%; Score 2810.5; DB 3; Length 991;  
 Best Local Similarity 93.6%; Pred. No. 2.6e-145;  
 Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;  
 Qy 1 SSDEEKNDVINNSQSIINTDDNQ-IKKEETNSDAIENRSKDIQTQSTTNVDENEATFLQK 60  
 Db 106 SSNEKNDVINNSQSIINTDDNQ-IKKEETNSDAIENRSKDIQTQSTTNVDENEATFLQK 164  
 Qy 61 TPQDNTHLTBEVKESSVSSNSIDTAQPSHTTINREESVQTSNDVDSHVSDFANS 120  
 Db 165 TPQDNTHLTBEVKESSVSSNSIDTAQPSHTTINSEASIQTSNEENSRSVDFANS 224  
 Qy 121 KIKESNTSGKEENTIEQPNKVKEDSTSPQSGVTNIDKISNODELNLNPINEYENKAR 180  
 Db 225 KIKESNTSGKEENTIEQPNKVKEDSTSPQSGVTNIDKISNODELNLNPINEYENKAR 284  
 Qy 181 PLSTTSAPSKRVTVNQLAAEQGSNNVHLIKVTDQSIETEGYDDSEGVIKAHAENLIYD 240  
 Db 285 PLSTTSAPSKRVTVNQLAAEQGSNNVHLIKVTDQSIETEGYDDSEGVIKAHAENLIYD 344  
 Qy 241 VTPEVDDKVKSGDGTMTVDIDKNTVPSDLTDSFTPIKIDNSGEIATGTVDNKNKQIYTT 300  
 Db 345 VTPEVDDKVKSGDGTMTVDIDKNTVPSDLTDSFAIPKIDNSGEIATGTVDNKNKQIYTT 404  
 Qy 301 FTDYVDKVENIKAHKLKTSYIDKSKVPNNNTKLDVEYKLTALSSVNTKITVEYQRPENRT 360  
 Db 405 FTDYVDKVENIKAHKLKTSYIDKSKVPNNNTKLDVEYKLTALSSVNTKITVEYQRPENRT 464  
 Qy 361 ANLOSMTNIDTKNHTVEQTIYINPLYSKATETNWNISGNGDEGTTIIDSTIIKVKVG 420  
 Db 465 ANLOSMTNIDTKNHTVEQTIYINPLYSKATETNWNISGNGDEGTTIIDSTIIKVKVG 524

QY 421 DNQNLPSNRNIYDYSEYEDVTNDYALQGNNDVNINFGNIDSPYIIKVISKYDPNKDDY 480  
 Db 525 DNQNLPSNRNIYDYSEYEDVTNDYALQGNNDVNINFGNIDSPYIIKVISKYDPNKDDY 584  
 QY 481 TTIOQTVMQTTINEYTGERTASYDNTIAFSTSSGOGGDLPPKTYIKIGDYVWEDVDK 540  
 Db 585 TTIOQTVMQTTINEYTGERTASYDNTIAFSTSSGOGGDLPPKTYIKIGDYVWEDVDK 644  
 QY 541 DGIQNTNDNEKPLSNVLVLTYPDGTGSKSVRTDEDEGKYQFDG 582  
 Db 645 DGIQNTNDNEKPLSNVLVLTYPDGTGSKSVRTDEDEGKYQFDG 686

RESULT 7  
 ABM79020  
 ID ABM79020 standard; protein; 549 AA.  
 XX  
 AC ABM79020;  
 DT 15-JAN-2004 (first entry)  
 XX  
 Staphylococcus epidermidis polypeptide.  
 DE  
 XX  
 Infection; antibacterial; vaccine.  
 KW  
 XX  
 Staphylococcus epidermidis.  
 OS  
 XX  
 WO2003076470-A1.  
 FN  
 XX  
 18-SEP-2003.  
 PD  
 XX  
 05-MAR-2003; 2003WO-US006415.  
 PF  
 XX  
 05-MAR-2002; 2002US-0361324P.  
 PR  
 XX  
 (INHI-) INHIBITEX INC.  
 PA  
 (TEXA ) UNIV TEXAS A & M SYSTEM.  
 XX  
 Patti JM, Hutchins JT, Hall A, Domanski P, Patel P, Hook M;  
 PI Robbins J, Vernachio J, Bowden MG;  
 PI  
 XX  
 WPI; 2003-722324/68.  
 DR  
 XX  
 New antibody recognizing a Staphylococcus epidermidis protein comprising  
 PT SdrG N1N2N3, SdrG N2N3 or SdrGR2 useful for preparing a composition for  
 PT treating or preventing a coagulase-negative staphylococcal infection.  
 XX  
 PS  
 Claim 20; Page 37; 78pp; English.  
 XX  
 The present sequence comprises the protein sequence of a polypeptide of a  
 CC coagulase-negative staphylococcus epidermidis. A claimed monoclonal  
 CC antibody recognises this protein and is used in a claimed method of  
 CC treating or preventing a coagulase-negative staphylococcal infection in a  
 CC human or animal, e.g. a nosocomial coagulase-negative staphylococcal  
 CC infection in low birth weight infants  
 XX  
 SQ Sequence 549 AA;

Query Match 88.6%; Score 2670; DB 7; Length 549;  
 Best Local Similarity 98.7%; Pred. NO. 5.7e-138;  
 Matches 519; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSDEKNDVINNNOSINTDDNNQIIKKEETNNYDIEKRSSEDRTESTNVNDEATFLOK 60  
 Db 24 SSDEKNDVINNNOSINTDDNNQIIKKEETNNYDIEKRSSEDRTESTNVNDEATFLOK 83  
 QY 61 TPQDNTHLTHEEVKESSVESNSSIDTAQPSHTTINRESVOTSDNVEDSHVSDFANS 120  
 Db 84 SPQDNTHLTHEEVKESSVESNSSIDTAQPSHTTINRESVOTSDNVEDSHVSDFANS 143  
 QY 121 KIKESNTESGKEENTIEOPNPKVEDSTTSQPSGYTNIDEKISNODELLNLPINEYENKAR 180  
 Db 144 KIKESNTESGKEENTIEOPNPKVEDSTTSQPSGYTNIDEKISNODELLNLPINEYENKAR 203

QY 181 PLSTTSAQPSIKRVTNQLAAEQGSNNVHLLIKVTDQSTTEGYDDSEGVIKAHDAENLIYD 240  
 Db 204 PLSTTSAQPSIKRVTNQLAAEQGSNNVHLLIKVTDQSTTEGYDDSEGVIKAHDAENLIYD 263  
 QY 241 VTFEYDDKVKSGDWTMTVDIDKNTVPSDLTDSFTTIPKIKDMSGELIATGTVDNKNKQIYTT 300  
 Db 264 VTFEYDDKVKSGDWTMTVDIDKNTVPSDLTDSFTTIPKIKDMSGELIATGTVDNKNKQIYTT 323  
 QY 301 FTDYVDKYENIKAHLLKLTYSIDKSKVPNNNTKLDVEYKLTALSSVNKTIITVEYQRPENRT 360  
 Db 324 FTDYVDKYENIKAHLLKLTYSIDKSKVPNNNTKLDVEYKLTALSSVNKTIITVEYQRPENRT 383  
 QY 361 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVNSGNDGEGSTIIDSTIIKVKYKVG 420  
 Db 384 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVNSGNDGEGSTIIDSTIIKVKYKVG 443  
 QY 421 DNQNLPSNRNIYDYSEYEDVTNDYALQGNNDVNINFGNIDSPYIIKVISKYDPNKDDY 480  
 Db 444 DNQNLPSNRNIYDYSEYEDVTNDYALQGNNDVNINFGNIDSPYIIKVISKYDPNKDDY 503  
 QY 481 TTIOQTVMQTTINEYTGERTASYDNTIAFSTSSGOGGDLPPK 526  
 Db 504 TTIOQTVMQTTINEYTGERTASYDNTIAFSTSSGOGGDLPPK 549

RESULT 8  
 ABM79015  
 ID ABM79015 standard; protein; 560 AA.  
 XX  
 AC ABM79015;  
 DT 15-JAN-2004 (first entry)  
 XX  
 Staphylococcus epidermidis SdrG N1N2N3 domain.  
 DE  
 XX  
 SdrG; surface protein; infection; antibacterial; vaccine.  
 KW  
 XX  
 Staphylococcus epidermidis.  
 OS  
 XX  
 WO2003076470-A1.  
 FN  
 XX  
 18-SEP-2003.  
 PD  
 XX  
 05-MAR-2003; 2003WO-US006415.  
 PF  
 XX  
 05-MAR-2002; 2002US-0361324P.  
 PR  
 XX  
 (INHI-) INHIBITEX INC.  
 PA (TEXA ) UNIV TEXAS A & M SYSTEM.  
 XX  
 Patti JM, Hutchins JT, Hall A, Domanski P, Patel P, Hook M;  
 PI Robbins J, Vernachio J, Bowden MG;  
 PI  
 XX  
 WPI; 2003-722324/68.  
 DR  
 N-PSDB; ACP80624.  
 DR  
 XX  
 New antibody recognizing a Staphylococcus epidermidis protein comprising  
 PT SdrG N1N2N3, SdrG N2N3 or SdrGR2 useful for preparing a composition for  
 PT treating or preventing a coagulase-negative staphylococcal infection.  
 XX  
 PS  
 Claim 27; Page 24-25; 78pp; English.  
 XX  
 The present sequence comprises the protein sequence of the N1N2N3 region  
 CC (amino acids 50-597), or putative A domain, of the SdrG surface protein  
 CC of coagulase-negative staphylococcus epidermidis. A claimed antibody  
 CC recognises a protein selected from SdrG N1N2N3, N2N3 and TR2. The  
 CC antibody may be a monoclonal antibody, including a chimeric, murine,  
 CC humanized, human or single chain monoclonal antibody, which prevents a  
 CC coagulase-negative staphylococcal infection in a human or animal by  
 CC inhibiting binding of staphylococcus bacteria to fibrinogen. Such  
 CC antibodies can be used to treat or prevent staphylococcal infections  
 CC including nosocomial coagulase-negative staphylococcal infections in low







Db 124 TETKPN--IEKNVPSDTKTATEDTSVILEEKKAPNNT--NNDVTTKPS-----TS 172

Qy 116 DFANSKIKESNT---ESGKEENTIEQPNKVKEDSTTSQPSGYTNI--DEKIS-NOBELNLNP 171

Db 173 EPSTSEIQTKPTTPOESTNIENSQPTSPKVD--NQVTDATPKPKNVSVKEELKNP 229

Qy 172 -----INEYENKARPLSTTSQAQSIKEV-----TUNOLAAEQGSNNVHLIKWTDQ 216

Db 230 EKLKELVRNDSNTDHTKTPVATPTSVAPKRVNAKMRFAVAQPAASNNVNDLIKVTQK 289

Qy 217 SITGYYDDEGIVIKAHDAENLIYDVTPEVDKVGSDTMTVDIDKNTVPSDLTDSFTIPK 276

Db 290 TIKVG-DGKDNVAAAHGDKDIEYDTEFTIDNKVKGGDTMTINYDKNVPISDLTDKNPID 348

Qy 277 IKONSGEIIATGYDNNKKNQITTYFTDYVDKYENIKAHKLKLTYSYDIXSKVPNNNTKLDVE 336

Db 349 ITDPSGEVIAGTTPDKATKQITTYFTDYVDKYEDIKSLRTLYSYIDKKTVP-NETSLNT 407

Qy 337 YKTALSSVNTKITVEYQRPNEKRTANLQSMFTNIDTKNHTVEQTIYINPLYSAKETNVN 396

Db 408 PATAGKETSQNVTVDYQDPWVHGDSNIQSIPTKLEDEKQIEQIYVNPPLKKSATNTKVD 467

Qy 397 ISGNG-----DEGSTIIDSTIIKVKYKVGDNQNLPSNRIYDYSEVEDVTND--DYA 446

Db 468 IAGSQVDYGNIKLNGSTIIDQNTKTEIKVKYKVSQQLPQSNRIYDYSQVEDVTISQFDNK 527

Qy 447 QLGNNVNDVNFNGNIDSPYIIKVKYKDPNKDDVTTIQQTVMOTTINEYTGEPRTASYD 506

Db 528 KSPFNNAVTLDFGDIINSAYIIKVKYSKVTPTSDGELDIAQGTSMRTT-DKY-GYNYAGYS 585

Qy 507 NTIAFSTSSGOGGLD-PPEKTYKIGDYWEDVDKDGIONTNDNEKELSNVLVTLTYPDG 565

Db 586 NFVITSNDTGGDGTVPPEELKXIGDYWEDVDKDGVGQGTDSKEKPMANVLVTLTYPDG 645

Qy 566 TSKSVRTDECKYQFDG 582

Db 646 TTKSVRTDANGHYEFGG 662

RESULT 12

ABJ18982

ID ABJ18982 standard; protein; 1166 AA.

XX

AC ABJ18982;

XX

XX

DT 06-MAR-2003 (first entry)

XX

DE Pathogen specific antigen related staphylococcal protein SEQ ID No 153.

XX Antibacterial; virucide; fungicide; protozoacide; cytostatic; anti-HIV;

KW hyperimmune; serum-reactive; antigen; pathogen; tumour; allergen;

KW auto-immunity; vaccine; staphylococcal infection; antibody; cancer;

KW autoimmune disease; HIV; hepatitis.

XX Staphylococcus sp.

OS

XX WO200259148-A2.

PN

XX

XX

PD 01-AUG-2002.

XX

PF 21-JAN-2002; 2002WO-EP0000546.

XX

PR 26-JAN-2001; 2001AT-00000130.

XX

XX (CIST-) CISTEM BIOTECHNOLOGIES GMBH.

PA

XX Meinke A, Nagy E, Von Ahean U, Klade C, Henics T, Zauner W;

PI Minh DB, Vytvytska O, Etz H, Dryla A, Weichhart T, Hafner M;

PI Tempelmaier B;

XX

XX WPI; 2003-075410/07.

XX

PT Identifying, isolating and producing hyperimmune serum-reactive antigens

PT from a pathogen, for preparing vaccine or medicament for treating or preventing e.g. staphylococcal infections, comprises providing antibody preparation.

PT

XX Example 7; Page 173; 252pp; English.

XX

CC The invention relates to a novel method for identifying, isolating and producing hyperimmune serum-reactive antigens from a pathogen, tumour, allergen, a tissue or host prone to auto-immunity, where the antigens are used in a vaccine, comprises providing antibody preparation from a plasma pool of a type of animal, or individual sera with antibodies against the specific pathogen, tumour, allergen, tissue or host prone to auto-immunity. The hyperimmune serum-reactive antigens comprising any of the 62 sequences of 53-2261 amino acids fully defined in the specification, or their hyperimmune fragments are useful for the manufacture of a pharmaceutical preparation, particularly a vaccine against staphylococcal infections or colonisation against *S. aureus* or *S. epidermidis*. The preparation of antibodies is useful for the manufacture of a medicament for treating or preventing staphylococcal infections or colonisation against *S. aureus* or *S. epidermidis*. The antibody preparations may also be used for diagnostic and imaging purposes. Other conditions that can be treated include cancer, autoimmune diseases or infections caused by viral (e.g. HIV, hepatitis A, B or C), fungal or protozoan pathogens. This sequence represents a staphylococcal protein relating to the method for identifying and producing pathogen specific antigens of the invention

CC

XX Sequence 1166 AA;

SQ

Query Match 37.9%; Score 1142.5; DB 6; Length 1166;

Best Local Similarity 42.7%; Pred. No. 4.8e-54;

Matches 262; Conservative 107; Mismatches 196; Indels 49; Gaps 18;

Qy 1 SSDEKNDVNNQSIINTDDNQI--IKKEETNNYDGIKRSSEDRSTTNNVDNEAFL 58

Db 66 ATTSNKEVVSETENNSTENNSTNPIKKE--TWDSQPEAKKESTSSQKQNNVTAT 123

Qy 59 QKTPQDNTHLTEREVKES-----SSVESNSSIDTAQPSHTTINREESVQTSNDVEDSHVS 115

Db 124 TETKPN--IEKNVPSDTKTATEDTSVILEEKKAPNNT--NNDVTTKPS-----TS 172

Qy 116 DFANSKIKESNT---ESGKEENTIEQPNKVKED---STTSQPSGYTNI--DEKISNOBEL 167

Db 173 EPSTSEIQTKPTTPOESTNIENSQPTSPKVDNQVTDATNPKEPNNVSKELKKNPEKL 232

Qy 168 LNLPISE--YENKARPLSTTSQAQSIKEV-----TVNOLAAEQGSNNVHLIKWTDQSI 219

Db 233 KELVRNDSNTDHTKTPVATPTSVAPKRVNAKMRFAVAQPAASNNVNDLIKVTQKTIK 292

Qy 220 EGYDDSEGVIKAHDAENLIYDVTPEVDKVGSDTMTVDIDKNTVPSDLTDSFTIPKID 279

Db 293 VGDGKDNVAAAHGDKDIEYDTEFTIDNKVKGGDTMTINYDKNVPISDLTDKNPIDID 351

Qy 280 NSGEIATGYDNNKKNQITTYFTDYVDKYENIKAHKLKLTYSYDIXSKVPNNNTKLDVEYKT 339

Db 352 PSGEVIAGTTPDKATKQITTYFTDYVDKYEDIKSLRTLYSYIDKKTVP-NETSLNLFAT 410

Qy 340 ALSSVNTKITVEYQRPNEKRTANLQSMFTNIDTKNHTVEQTIYINPLYSAKETNVNISG 399

Db 411 AGKETSQNVTVDYQDPWVHGDSNIQSIPTKLEDEKQIEQIYVNPPLKKSATNTKVDIAG 470

Qy 400 NG-----DEGSTIIDSTIIKVKYKVGDNQNLPSNRIYDYSEVEDVTND--DYAQLG 449

Db 471 SQVDDYGNIKLNGSTIIDQNTKTEIKVKYKVSQQLPQSNRIYDYSQVEDVTISQFDNK 530

Qy 450 NNDVNDVNFNGNIDSPYIIKVKYKDPNKDDVTTIQQTVMOTTINEYTGEPRTASYDNTI 509

Db 531 SNNVATLDFGDIINSAYIIKVKYSKVTPTSDGELDIAQGTSMRTT-DKY-GYNYAGYSNFI 588

Qy 510 AFSTSSGOGGLD-PPEKTYKIGDYWEDVDKDGIONTNDNEKELSNVLVTLTYPDGTSK 568

Db 589 VTSNDTGGDGTVPPEELKXIGDYWEDVDKDGVGQGTDSKEKPMANVLVTLTYPDGTTK 648

Qy 569 SVRTDECKYQFDG 582

||||| : | : | : |  
649 SVRTDANGHYBFGG 662

Db  
RESULT 13  
ABU42327  
ID ABU42327 standard; protein; 1141 AA.  
XX AC ABU42327;  
XX 19-JUN-2003 (first entry)  
DT Protein encoded by Prokaryotic essential gene #27854.  
XX DE Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX KW Staphylococcus aureus.  
XX OS WO200277183-A2.  
XX PN 03-OCT-2002.  
XX PD 21-MAR-2002; 2002WO-US009107.  
XX PF 21-MAR-2001; 2001US-00815242.  
XX PR 06-SEP-2001; 2001US-00948993.  
XX PR 25-OCT-2001; 2001US-0342923P.  
XX PR 08-FEB-2002; 2002US-00072851.  
XX PR 06-MAR-2002; 2002US-0362699P.  
XX PA (ELIT-) ELITRA PHARM INC.  
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX WPI; 2003-029926/02.  
XX N-PSDB; ACA46197.  
XX DR New antisense nucleic acids, useful for identifying proteins or screening  
XX PT for homologous nucleic acids required for cellular proliferation to  
XX PT isolate candidate molecules for rational drug discovery programs.  
XX PS Claim 25; SEQ ID NO 70251; 1766pp; English.  
XX CC The invention relates to an isolated nucleic acid comprising any one of  
XX CC the 6213 antisense sequences given in the specification where expression  
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid  
XX CC encoding a polypeptide whose expression is inhibited by the antisense  
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
XX CC polypeptide or its fragment whose expression is inhibited by the  
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding  
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
XX CC proliferation or the activity of a gene in an operon required for  
XX CC proliferation; (7) identifying a compound that influences the activity of  
XX CC the gene product or that has an activity against a biological pathway  
XX CC required for proliferation, or that inhibits cellular proliferation; (8)  
XX CC identifying a gene required for cellular proliferation or the biological  
XX CC pathway in which a proliferation-required gene or its gene product lies  
XX CC or a gene on which the test compound that inhibits proliferation of an  
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
XX CC compound's activity; (11) a culture comprising strains in which the gene  
XX CC product is overexpressed or underexpressed; (12) determining the extent  
XX CC to which each of the strains is present in a culture or collection of  
XX CC strains; or (13) identifying the target of a compound that inhibits the  
XX CC proliferation of an organism. The antisense nucleic acids are useful for  
XX CC identifying proteins or screening for homologous nucleic acids required  
XX CC for cellular proliferation to isolate candidate molecules for rational  
XX CC drug discovery programs, or for screening homologous nucleic acids  
XX CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
XX CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
XX CC the target prokaryotic essential genes. Note: The sequence data for this  
XX CC patent did not form part of the printed specification, but was obtained

CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 1141 AA;  
Query Match 37.8%; Score 1137.5; DB 6; Length 1141;  
Best Local Similarity 43.3%; Pred. No. 8.8e-54;  
Matches 266; Conservative 97; Mismatches 193; Indels 59; Gaps 19;  
Qy 1 SSDEKNDVINNQSINTDDNNQIIKKEETNNVDGIEKRSRDETESTNVDENATFLOK 60  
Db 69 SDNKEVSVETENNSTTENDSTNPI--KKEETNT--DSOPEAKEESTTSQQQNNVTATTE 125  
Qy 61 TPQONTHLTREEVKES---SSVESSNSIDTAQOPSHHTINREESVCTSDNVEDSHVSDF 117  
Db 126 TKQN--IEKNVKPSTDKTATEDTSVILEKAPNYT--NNDVTTKPS----- 170  
Qy 118 ANSKIKESNT---ESGKEENTIEQPNKVKEDSTTSQPSGYNTIDEKIS--NDELLNLP-- 171  
Db 171 -TSEIQTKPTTPOESTNIENSQPQTPSKVD---NQVTDATNPKEPVNVSKELKNPEK 226  
Qy 172 -----INEYENKARPLST--TSAQP-----SIKRVTVNQAAEQGSNVNHLIKVTDQSI 218  
Db 227 LKELVRNDNTDRSTKPVATAPT--SVAPKRLNAXKRFAPAQAASNNVNDLITVTKQTI 286  
Qy 219 TEGYDDSEGVIKAHDAENLIYDVTFEVDKVGSDTMTVDIDKNTVPSDLTDSFTIPKIK 278  
Db 287 KVG-DGKDNVAAADGKDIEYDTEFTIDNKVKGDFTWINDKNTVPSDLTKNDPIDIT 345  
Qy 279 DNSGEIATGTYDNKNKQIYTTFTDVKYKENTKAHLKLTYSIDKSKVPNNNTKLDVEYK 338  
Db 346 DPSGEVIAKGTDFKATKQIYTTFTDVKYKEDIKARLTLYSIDKQAVP--NETSLNLTEA 404  
Qy 339 TALSSVNNKTTTVEYQRPENRTANLOSMTNIDTKNHTVEOTIYNPLRYSAKETNNNIS 398  
Db 405 TAGKETSONVSDIYQDPMVHGDSNIQSIIFKLDENKQIEQQIYVAPLKKATNTNKVDIA 464  
Qy 399 GNG-----DEGSIIDDSIIIVKYKVGDNQNLPSNRIYDYSEYEDVTND--DYAQL 448  
Db 465 GSQVDDYGNIKLNGSTIIDQNTETIKVKYKVPNQQLPQSNRIYDFSQYEDVTSQDFNKS 524  
Qy 449 GNNNDVINFGNIDSPYIIKISKYDPNKDDYTTIQTVMQTTINEYTGFEFTASYDNT 508  
Db 525 FSNNVATLDFGDNINSAYIIKXWSKYTPTSDGELDIAQGTSMRTT-DKY-GYVNYAGYSNF 592  
Qy 509 IAPSTSSGQOQDL-DPEKTYKIGDYVWEDVDKGTQNTNDNEKPLSNVLATLTYPDGTS 567  
Db 583 IVTSNDTGGDGTVPKEEKLYKIGDYVWEDVDKGTQNTNDNEKPLSNVLATLTYPDGTT 642  
Qy 568 KSVRTDEDDGKYQPDG 582  
Db 643 KSVRTDANGHYBFGG 657  
RESULT 14  
ABU42520  
ID ABU42520 standard; protein; 670 AA.  
XX AC ABU42520;  
XX 19-JUN-2003 (first entry)  
DT Protein encoded by Prokaryotic essential gene #28047.  
XX DE Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX KW Staphylococcus epidermidis.  
XX OS WO200277183-A2.  
XX PN 03-OCT-2002.  
XX PD 21-MAR-2002; 2002WO-US009107.  
XX PF

XX 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX (ELIT-) ELITRA PHARM INC.  
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX  
XX WPI; 2003-029926/02.  
DR N-PSDB; ACA46390.  
XX  
XX New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
XX Claim 25; SEQ ID NO 70444; 1766pp; English.  
XX  
XX The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: the sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 670 AA;  
Query Match 37.1%; Score 1116.5; DB 6; Length 670;  
Best Local Similarity 98.2%; Pred. No. 6.1e-53;  
Matches 213; Conservative 0; Mismatches 1; Indels 3; Gaps 1;  
Qy 366 MFTNIDKNTVEQTIYNPLRYSAKETNNVNSGNDEGSTIIDSTTIKVKYKVDNQNL 425  
Db 1 MFTNIDKNTVEQTIYNPLRYSAKETNNVNSGNDEGSTIIDSTTIKVKYKVDNQNL 60  
Qy 426 PDSNRIYDSEYEDVNDVIAQLGNNDVNFNFGNIDSPYIIKVIKSYDPNKKDYTTIQQ 485  
Db 61 PDSNRIYDSEYEDVNDVIAQLGNNDVNFNFGNIDSPYIIKVIKSYDPNKKDYTTIQQ 120  
Qy 486 TVTWTQTINETYGFRFASVDNTIAFTSSGQGGDLPPKTYKIGYVWEDVDKGIQN 545  
Db 121 TVTWTQTINETYGFRFASVDNTIAFTSSGQGGDLPPKTYKIGYVWEDVDKGIQN 177  
Qy 546 TNDNEKPLSNVLVTLPDGTSKSVRTDEDEKGYQFDG 582  
|||||

Db 178 TNDNEKPLSNVLVTLPDGTSKSVRTDEDEKGYQFDG 214  
RESULT 15  
AAG82803  
ID AAG82803 standard; protein; 278 AA.  
XX AAG82803;  
XX  
XX 03-SEP-2001 (first entry)  
XX  
XX S. epidermidis open reading frame protein sequence SEQ ID NO:2700.  
XX Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;  
KW endocarditis.  
XX Staphylococcus epidermidis.  
XX  
XX WO200134809-A2.  
XX  
XX 17-MAY-2001.  
XX  
XX 09-NOV-2000; 2000WO-US030782.  
XX  
XX 09-NOV-1999; 99US-0164258P.  
XX (GLAX) GLAXO GROUP LTD.  
XX  
XX Kimmerly WJ;  
XX  
XX WPI; 2001-316495/33.  
DR N-PSDB; AAH53653.  
XX  
XX Nucleic acids encoding polypeptides from *Staphylococcus epidermidis*,  
PT useful for vaccinating against infections, e.g. endocarditis.  
XX  
XX Claim 18; Page 705; 2188pp; English.  
XX  
XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides  
CC (II), given in AAG81454 to AAG83120, from *Staphylococcus epidermidis*. (I)  
CC and (II) can have antibacterial activity and therefore can be used in  
CC vaccination. The nucleic acids (I) may be used to produce the S.  
CC epidermidis polypeptides (II) via the production of vectors containing  
CC them which are used to produce hosts cells which express the  
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be  
CC used to vaccinate subjects and to raise antibodies against the bacteria.  
CC The polypeptides may also be used to assay for other inhibitors of their  
CC activity and therefore identify compounds that may be used for the  
CC treatment of *S. epidermidis* infections, e.g. endocarditis. AAH53971 to  
CC AAH5090 represent specifically claimed *S. epidermidis* genomic DNA  
CC polynucleotide sequences from the present invention. AAH5091 to AAH5098  
CC represent oligonucleotide sequences and primers which are used in the  
CC exemplification of the present invention. N.B. The present invention  
CC specifically claims all the polynucleotide sequences given in the  
CC sequence listing of the present specification, however the sequence  
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given  
CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present  
CC for SEQ ID NO:4455 to 4464  
XX  
XX Sequence 278 AA;  
Query Match 28.4%; Score 857; DB 4; Length 278;  
Best Local Similarity 93.0%; Pred. No. 2.9e-39;  
Matches 172; Conservative 3; Mismatches 8; Indels 2; Gaps 1;  
Qy 1 SSDEKNDVNNQSNINTDNNQIIKKEETNNYDGIKRSSEDRTESTTNVDENATFLQK 60  
Db 75 SSDEKNDVNNQSNINTDNNQIIKKEETNNYDGIKRSSEDRTESTTNVDENATFLQK 134  
Qy 61 TPQDNTHLTETEEVKESSESSVESSSSIDTAQPSHTTINREESVQTSNVDSDSHVDFANS 120  
Db 135 SPQDNTHLTETEEVKESSESSVESSSSIDTAQPSHTTINREESVQTSNVDSDSHVDFANS 194  
|||||

QY 121 KIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDEKISNQDELLNLPINEYENKAR 180  
Db |||||  
195 KIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDEKISNQDELLNLPINEY--KVT 252  
QY 181 PLSTT 185  
Db |||  
253 KLSVT 257

Search completed: October 5, 2004, 20:48:05  
Job time : 70.362 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:44:57 ; Search time 17.0358 Seconds  
(without alignments)  
1763.712 Million cell updates/sec

Title: US-10-806-288-13  
Perfect score: 3013  
Sequence: 1 SSDEEKNDVINNQSIINTDD.....PDGTSKSVRTDECKYQFDG 582

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2810.5	93.3	930	4	US-09-134-001C-5314
2	2810.5	93.3	930	4	US-09-386-962C-10
3	1144.5	38.0	1166	4	US-09-200-650E-7
4	576	19.1	1315	4	US-09-200-650E-5
5	499	16.6	1742	4	US-09-386-962C-4
6	455	15.1	930	4	US-09-200-650E-3
7	454	15.1	936	4	US-08-956-171E-5249
8	452	15.0	933	3	US-08-293-728-2
9	452	15.0	933	3	US-09-421-868-2
10	396	13.1	918	4	US-09-200-650E-1
11	349	11.6	345	3	US-08-856-253-7
12	300.5	10.0	1027	4	US-08-956-171E-5254
13	259.5	8.6	1112	2	US-08-714-402-2
14	259.5	8.6	1161	4	US-09-327-536-2
15	220.5	7.3	3696	4	US-09-134-001C-5080
16	209	6.9	630	3	US-08-973-462-9
17	208.5	6.9	251	4	US-08-956-171E-5252
18	203.5	6.8	2504	4	US-09-328-352-5821
19	202.5	6.7	368	4	US-09-134-000C-4597
20	202	6.7	676	4	US-09-134-001C-4318
21	201.5	6.7	1177	4	US-09-134-001C-5106
22	197	6.5	2314	4	US-09-268-347-49
23	195	6.5	1335	4	US-09-134-001C-3716
24	195	6.5	10182	4	US-09-134-001C-3159
25	192.5	6.4	1306	4	US-09-134-000C-6670
26	192.5	6.4	1786	3	US-08-973-462-8
27	190.5	6.3	1833	4	US-08-621-944A-4

28	190.5	6.3	1833	4	US-08-945-567D-4	Sequence 4, Appli
29	190.5	6.3	1992	4	US-08-621-944A-3	Sequence 3, Appli
30	190.5	6.3	1992	4	US-08-945-567D-3	Sequence 3, Appli
31	189.5	6.3	1664	1	US-09-599-652-2	Sequence 2, Appli
32	189.5	6.3	1664	2	US-08-642-846-2	Sequence 2, Appli
33	189.5	6.3	1664	4	US-09-264-604-2	Sequence 2, Appli
34	188.5	6.3	1435	2	US-08-568-459A-4	Sequence 4, Appli
35	188.5	6.3	1435	2	US-08-487-826B-4	Sequence 4, Appli
36	188.5	6.3	1435	4	US-09-210-288-4	Sequence 4, Appli
37	187	6.2	1416	4	US-09-071-035-404	Sequence 404, App
38	187	6.2	1448	4	US-09-071-035-402	Sequence 8, Appli
39	186.5	6.2	3088	4	US-09-562-702A-8	Sequence 4, Appli
40	186.5	6.2	3089	4	US-09-562-702A-4	Sequence 2, Appli
41	186.5	6.2	3110	4	US-09-562-702A-2	Sequence 2, Appli
42	186.5	6.2	3110	4	US-09-562-702A-6	Sequence 6, Appli
43	186.5	6.2	3110	4	US-09-561-709B-7	Sequence 7, Appli
44	186.5	6.2	3111	2	US-08-460-309-4	Sequence 4, Appli
45	186.5	6.2	3111	2	US-08-125-077-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1  
US-09-134-001C-5314  
; Sequence 5314, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; FILE REFERENCE: GPC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 5314  
; LENGTH: 930  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-5314

Query Match	93.3%	Score 2810.5;	DB 4;	Length 930;
Best Local Similarity	93.6%	Pred. No. 2e-158;		
Matches	545;	Conservative	15;	Mismatches 21;
			Indels	1;
			Gaps	1;
Qy	1	SSDEEKNDVINNQSIINTDDNNQIIKKEETNNVDGIEKRSIEDRTSTNTVDENEATFLOK	60	
Db	73	SSNEEKNDVINNQSINTDDNQ-IKKEETNSDAIENRSKDIQTSTNTVDENEATFLOK	131	
Qy	61	TPQDNTHLTEEVKESSVSSNSSIDTAOPSHTTINREESVOTSDNVEDSHVSDFANS	120	
Db	132	TPQDNTLKEEVKPESSVSSNSSMDTAQPSHTTINSEASIQTSNEENSRVSDFANS	191	
Qy	121	KIKESNTESKEENTIEQPNKVEDSTTSQPSGYNTIDKISNQDELLNLPINEYENKAR	180	
Db	192	KIESNTESKEENTIEQPNKVEDSITSQPSGYNTIDKISNQDELLNLPINEYENKVR	251	
Qy	181	PLSTTSAPQISKRVTNNQLAEOGSNNVHLIKVTDQSITEGYDDSEGVKAHAENLIYD	240	
Db	252	PLSTTSAPQSKRVTNNQLAEOGSNNVHLIKVTDQSITEGYDDSDGIKAHAENLIYD	311	
Qy	241	VTFEVDKVKSGTMTVDIDKNTVPSDLTDSFTPKIKDNGSEIATGTVDNKNQIITYT	300	
Db	312	VTFEVDKVKSGTMTVNDKNTVPSDLTDSFALPKIKDNGSEIATGTVDNKNQIITYT	371	
Qy	301	FTDVVDKYNKIAHLKLTYSIDKSKVNNNTKLDVEYKLTALSSVKNKTIITVEYQRPENRT	360	
Db	372	FTDVVDKYNKIAHLKLTYSIDKSKVNNNTKLDVEYKLTALSSVKNKTIITVEYQRPENRT	431	

QY 361 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVMISGNDEGSGTIIDSTIIKVKYVG 420  
Db 432 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVMISGNDEGSGTIIDSTIIKVKYVG 491  
QY 421 DNQNLPSNRIYDYSEYEDVTNDYAOQLGNNDVNIINFNGIDSPYIIKVISKYPDNKODY 480  
Db 492 DNQNLPSNRIYDYSEYEDVTNDYAOQLGNNDVNIINFNGIDSPYIIKVISKYPDNKODY 551  
QY 481 TTIQQTVMQTTINEYTGEBFRASVDNTIAFSTSSGQGGDLPPPEKTYKIGDYVWEDVK 540  
Db 552 TTIQQTVMQTTINEYTGEBFRASVDNTIAFSTSSGQGGDLPPPEKTYKIGDYVWEDVK 611  
QY 541 DGIQNTNDNEKPLSNVLVTLTPDGTGSKSVRTDEGKYQFDG 582  
Db 612 DGIQNTNDNEKPLSNVLVTLTPDGTGSKSVRTDEGKYQFDG 653

## RESULT 2

US-09-386-962C-10  
; Sequence 10, Application US/09386962C  
; Patent No. 6635473  
; GENERAL INFORMATION:  
; APPLICANT: FOSTER, Timothy  
; TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLO  
; FILE REFERENCE: P06335US2/BAS  
; CURRENT APPLICATION NUMBER: US/09/386,962C  
; CURRENT FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: 60/098,443  
; PRIOR FILING DATE: 1998-08-31  
; PRIOR APPLICATION NUMBER: 60/117,119  
; PRIOR FILING DATE: 1999-01-25  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 10  
; LENGTH: 930  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-386-962C-10

Query Match 93.3%; Score 2810.5; DB 4; Length 930;  
Best Local Similarity 93.6%; Pred. No. 2e-158;  
Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;  
QY 1 SSDEKNDVINNQSDINTDDNQIIKKEETNNYDGIKRSDBRTESTTNVDENEAFLQK 60  
Db 73 SSNEEKNDVINNQSDINTDDNQIIKKEETNSDAIENRSKDIQSTTNVDENEAFLQK 131  
QY 61 TPQDNTLHTEERVKESSSVESNSSIDTAQOPSHHTINREESVQTSNVEDSHVDFANS 120  
Db 132 TPQDNTLHTEERVKESSSVESNSSMDTAQOPSHHTINSEASIQTSNEENSRSVDFANS 191  
QY 121 KIKESNTESGKEENTIEQPNVKEDSTTSQPSGYTNIIDEKISNQDELINLPINEYENKAR 180  
Db 192 KIKESNTESGKEENTIEQPNKVEDSITSQPSYKNIDEKISNQDELINLPINEYENKVR 251  
QY 181 PLSTTSAPSPSKRVTVNQLAAEQSGNVNHLIKVTDQSITGEYDSDSEGVKAHAENLIYD 240  
Db 252 PLSTTSAPSPSKRVTVNQLAAEQSGNVNHLIKVTDQSITGEYDSDGIKAHAENLIYD 311  
QY 241 VTFEVDDKVGSGDTMTVDIDKNTVPSDLTSDFTIPKIKNSGELIATGVDNKNKIITYT 300  
Db 312 VTFEVDDKVGSGDTMTVNIIDKNTVPSDLTSDFAIPKIKNSGELIATGVDNTNKNKIITYT 371  
QY 301 FTDYVDKYENIKAHKLKTSYIDSKVPNNNTKLDVEYKNTALSSVKNKTIITVEYQRPENRT 360  
Db 372 FTDYVDKYENIKAHKLKTSYIDSKVPNNNTKLDVEYKNTALSSVKNKTIITVEYQRPENRT 431  
QY 361 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVMISGNDEGSGTIIDSTIIKVKYVG 420  
Db 432 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVMISGNDEGSGTIIDSTIIKVKYVG 491  
QY 421 DNQNLPSNRIYDYSEYEDVTNDYAOQLGNNDVNIINFNGIDSPYIIKVISKYPDNKODY 480

Db 492 DNQNLPSNRIYDYSEYEDVTNDYAOQLGNNDVNIINFNGIDSPYIIKVISKYPDNKODY 551  
QY 481 TTIQQTVMQTTINEYTGEBFRASVDNTIAFSTSSGQGGDLPPPEKTYKIGDYVWEDVK 540  
Db 552 TTIQQTVMQTTINEYTGEBFRASVDNTIAFSTSSGQGGDLPPPEKTYKIGDYVWEDVK 611  
QY 541 DGIQNTNDNEKPLSNVLVTLTPDGTGSKSVRTDEGKYQFDG 582  
Db 612 DGIQNTNDNEKPLSNVLVTLTPDGTGSKSVRTDEGKYQFDG 653

## RESULT 3

US-09-200-650E-7  
; Sequence 7, Application US/09200650E  
; Patent No. 6680195  
; GENERAL INFORMATION:  
; APPLICANT: Patti, Joseph M.  
; APPLICANT: Foster, Timothy J.  
; APPLICANT: Hook, Magnus A.O.  
; APPLICANT: Eichmann, Deirdre Ni  
; APPLICANT: Perkins, Samuel L.  
; TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aureus  
; FILE REFERENCE: P06283US2/BAS  
; CURRENT APPLICATION NUMBER: US/09/200,650E  
; CURRENT FILING DATE: 1998-11-25  
; PRIOR APPLICATION NUMBER: 60/066,815  
; PRIOR FILING DATE: 1997-11-26  
; PRIOR APPLICATION NUMBER: 60/098,427  
; PRIOR FILING DATE: 1998-08-31  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 1166  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-09-200-650E-7

Query Match 38.0%; Score 1144.5; DB 4; Length 1166;  
Best Local Similarity 42.6%; Pred. No. 7.9e-60;  
Matches 263; Conservative 108; Mismatches 191; Indels 55; Gaps 18;  
QY 1 SSDEKNDVINNQSDINTDDNQIIKKEETNNYDGIKRSDBRTESTTNVDENEAFL 58  
Db 66 ATTSNKEVSEVETENNSTTENNSTNPIKKE--TNTDSQPEAKKESSTSSQKQNNVTAT 123  
QY 59 QKTPQDNTLHTEERVKESSSVESNSSIDTAQOPSHHTINREESVQTSNVEDSHV 115  
Db 124 TETKPN--IEKNVPSPTDKTATEDTSVILEKAPNNT--NNDVTTKPS-----TS 172  
QY 116 DFANSKIKESNT---ESGKEENTIEQPNVKEDSTTSQPSGYTNIIDEKIS-NODELINLP 171  
Db 173 EPSTSEIQTKPTTPQSTNIEANSQOPTPSKVD--NQVTDATNPKEPVNVSKELKNP 229  
QY 172 -----INVEYENKARPLSTTSAPSPSKRV-----TVNQLAAEQSGNVNHLIKVTDQ 216  
Db 230 EKLKELVRNDSNTDHTKPVATAPTSVAPKRVNAKMFAPAFAAVASNNVNDLIKVTQ 289  
QY 217 SITEGVDDSEGVTKAHAENLIYDVTFEVDKVGSGDTMTVDIDKNTVPSDLTSDFTIPK 276  
Db 290 TIKVG-DGKNVNAAGDKDIEYDTEFTIDNKKVKGDTMTINIDKRVISDLTDDKNDDID 348  
QY 277 IKDNGEIIATGYDNKNKQIITYTFTDYVDKYENIKAHKLKTSYIDSKVPNNNTKLDVE 336  
Db 349 ITDPSGEVIAGKTFDKATKQITVTFTDYVDKYEDIKSRLLTLYSIDKTVP-NETSLNLT 407  
QY 337 YKTALSSVKNKTIITVEYQRPENRTALOSMFPNIDTKNHTVEQTIYINPLRYSAKETN 396  
Db 408 FATAGKETSNVTVDYQDPMVHGDSNIQSIIFTKIDEDKQTIIEQIYVNPFLKKSATNTKVD 467  
QY 397 ISGNG-----DRGSTIIDSTIIKVKYVGDNQLPDSNRIYDYSEYEDVTND-DYA 446  
Db 468 IAGSQVDDYGNIKLNGSGTIIDQNTIEKVKYKNSDQQLPQSNRIYDPSQYEDVTSQFDNK 527







Qy	121	KIKESNTSEGEENTIEQPNKVKEDSTTSQPSGYTNIDEKISNODELLN-LPINEYENKA	179
Db	153	--QTSNETTSDNTV-----SSVNSQNSTNAENVSTQDTSTTEATPSN---NES	198
Qy	180	RPLSTTSAQPSIKRVTVN-----QLAABQGSNVNHLIKVTDQ---SITEGYDDSGEV	228
Db	199	APQSTDSANKDVNQAVNTSAPRMAFSLAAVAADAPVAGTDITNQLTNVTVGI--DSGTT	257
Qy	229	IKAHDAENLIYDVTFEVDKVKYSGDGMTVDIDKNTVPSDLTSTFTPIKIKONSGRIIATG	288
Db	258	VYFHOAGVKULNYGFSVPNSAVKGDTFKITVPKELNLGVSTSTAKVPIIMAGD-QVLANG	316
Qy	289	TYDNKNKQITVTFDYVDKYENIKAKHLKLTGYIDSKSVPNNTKU-DVEYKFTALSS--VN	345
Db	317	VIDS-DGNVIYTFTDYVNTKDDVRATLTPAYID----PENVKKTGNVTLATIGISTTAN	371
Qy	346	KTITVEYORPNRENTANLOSMFTNIDTKNHTVBQTIYINPL--RYSAKETNVNISGNGDE	403
Db	372	KTVLVDYKYGKFNLSIKGTHIDKTNNTVRTIYVNPSPGDNVIA PVLTLGNLPNTDS	431
Qy	404	GSTIIDBSTIIKVKYKGDQNLPDSNRIDYSEYEDVTDNDYQAQGNNDNVNIFG----	459
Db	432	NALIDQOQNTSIKYKVDNAADLSSESYFV-NPENFEDVTNSVNIITPPNPQYKVEFNTPDD	490
Qy	460	NIDSPYIIKVISKYDPN-KDQVTTIQQVTVWQTTINEYTG-----PRTASVDNTIAFSTSG	516
Db	491	QITTPYIUVVNGHIDPNSKGD-----LALRSTLYGYNSIIWMSMWDNEVAFNGSG	543
Qy	517	QGOQ-DLP--PEKTYKIGDY--VMEVDVKD-GIQNTNDN	549
Db	544	SGDGIDKVPVPEQDEPGEIPIBPDSDSRPGSDSGSDS	582

## RESULT 8

```

US-08-293-728-2
; Sequence 2, Application US/08293728D
; Patent No. 6008341
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; APPLICANT: McDevitt, Damien L.
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
; FILE REFERENCE: 0344.105011
; CURRENT APPLICATION NUMBER: US/08/293.728D
; CURRENT FILING DATE: 1994-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 933
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-293-728-2

```

Query Match 15.0%; Score 452; DB 3; Length 933;

[illegible]



Db 84 TAKUPPIMAGD-QVLANGVIDS-DGNVIYTFDVTNWKDDVKALTMPAYID-----PENV 137  
Qy 331 TKL-DVEYKTALSS--VNKTIIVYORPNENRTANLQSMFTNIDTKNHTVEQTIYINPL- 386  
Db 138 KKTGNVLTATGIGSTANKTANLVLDYKYGKFPYVLSIKGTIDQIDKTNNTYRQTIYVPSG 197  
Qy 387 -RYSAKETNVMISNGDEGSIIDSTIIKYYKVGDNQNLPSNRRIYDSEYEDVTNDY 445  
Db 198 DNVIAPVLTKLTKPTNSALIDQNTSIKYKYVDNAADLSESYFV-NPENFEDVTNSVN 256  
Qy 446 AOLGNNDVNINFG---NIDSPYIIKISKYDPN-KDDYTTIQOTVTMOTTINEXYTG- 499  
Db 257 ITFPNPKYKVFYTFPDQITPTPIVVVNGHIDPNSKGD-----LALRSTLYGYSNI 309  
Qy 500 -FRTASYDNTIAFTSSGQGG-DLP--PEK 526  
Db 310 IWRSMWNEVAFNNGSGGIDKPVVPEQ 340

## RESULT 12

US-08-956-171E-5254  
; Sequence 5254, Application US/08956171E  
; Patent No. 6593114

## GENERAL INFORMATION:

APPLICANT: Charles Kunsch

Gil H. Choi

Patrick S. Dillon

Craig A. Rosen

Steven C. Barash

Michael R. Pannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5256

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/956.171E

FILING DATE: 20-Oct-1997

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/009,861

FILING DATE: January 5, 1996

APPLICATION NUMBER: 08/781,986

FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Mark J. Hyman

REGISTRATION NUMBER: 46,789

REFERENCE/DOCKET NUMBER: PB248P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (240) 314-1224

TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 5254:

SEQUENCE CHARACTERISTICS:

LENGTH: 1027 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 5254:

US-08-956-171E-5254

Query Match 10.0%; Score 300.5; DB 4; Length 1027;

Best Local Similarity 21.2%; Pred.No.5.1e-10;

Matches 123; Conservative 118; Mismatches 243; Indels 97; Gaps 25;

Qy 35 GIEKRSBDRTESTNTVDENEATFLQKTPQN-----THLTEEVEKSESSVSSSSSIDTAQ 90  
Db 39 GQDKEAAASEQKTTTVEEN-----GNSATDNKTSSETOTTATNVNHIETQSYNATV--TE 91  
Qy 91 QPSHTT--INFEER---SVQSDNVEDSHVDFANSKIKESNTESGKEENTLEQPNKVKEDS 146  
Db 92 QPSNATQVTTTEAPKVAQAPQTAQANITVKEBEVKE-----EAKPQVKETT 139  
Qy 147 TTSOPSG---YTNIDEKISNQDELLNLPINEYENKARPLSTTSAPSIKRVT-----V 196  
Db 140 QSQDNGSDQVLDLTPKATQNVAEQV---EVAQPTASESKPRVTRSDVAEAKA 195  
Qy 197 NQLAABEGSNVNHILKVTQDSITEGYDDSEGVKAHDAENLIYDVTFEVDDKVKSGDTMT 256  
Db 196 SNAKVETGTDVTSKVTVEIGSI-EGHNNTNKV-EPHAGORAVLYKLFKFNGLHQGDYFD 253  
Qy 257 VDIDKNTVPSDLTDSFTPIPKIDNS-----GEIIATGTYDNKNKOITVPTDYVDKYENI 311  
Db 254 FTLSNNVTHGVTARKVPEIKNGSVVMATGEVLEGG-----KIRYFTFNDIEDKVDV 306  
Qy 312 KAHLKLTSYIDKSKVPNNNTKLDVEYKLTALSSVN-----KTIITVEYQRPNERRTANLQSM 366  
Db 307 TABEINLFDPKTVQTNNGN-----QTITSTLNEEQTSKELDVKKYKDGIGNYVANLGS 360  
Qy 367 FTNIDTKQHTVEQTIYINPLRYSAKETNVMISNGDEGSIIDSTIIKVKV-VGDQNQL 425  
Db 361 IETFNKANRRFHVAFIKP--NNGKTTSTVTGTLMKGSQNGNQKQKVRIFEYLGNNEDI 418  
Qy 426 PDSNRIV---DYSEVEDVTNDYQAOLG--NNNDVINFGNIDSPYIIKVISKYDPNKO- 478  
Db 419 AKS--VYANTTDTSKFEVTSNMGSLNQLNGQSYSLNIENLDKTYVYVYDGEYLNLTDE 476  
Qy 479 -DYTTIQOTVTMOTTINEYTFGEFTASYDNTIAFTSSGQGGDLPP-----EKTYKIGD 532  
Db 477 VDFRT-QMVGHPEQLYKYYYDRGYTLTDWNLGLVLYSNKANGNEKNGPIIQNNKFEYK--- 532  
Qy 533 YVMEDVDKGIQNTNDNEKPLSNVLTLTVPDGTGSKSVRTD 573  
Db 533 ---EDTIKETLTGQYDKN-----LVTTVBEEYDSDSTLDD 564

## RESULT 13

US-08-714-402-2

; Sequence 2, Application US/08714402

; Patent No. 5910441

GENERAL INFORMATION:

APPLICANT: ROCHA, Claudia

APPLICANT: FISCHETTI, Vincent A.

TITLE OF INVENTION: FIBRONECTIN AND FIBRINOGEN BINDING

TITLE OF INVENTION: PROTEIN FROM GROUP A STREPTOCOCCI

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: BURNS, DOANE, SWECKER &amp; MATHIS

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/714,402

FILING DATE: 16-SEP-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300

REFERENCE/DOCKET NUMBER: 016921-097

TELECOMMUNICATION INFORMATION:

```

; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5080
; LENGTH: 3696
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5080
Query Match 7.3%; Score 220.5; DB 4; Length 3696;

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:49:18 ; Search time 68.491 Seconds  
(without alignments)  
2734.476 Million cell updates/sec

Title: US-10-806-288-13

Perfect score: 3013

Sequence: 1 SSDEEKNDVNNQNSINTDD.....PDGTSKSVRTDEGKVFQFDG 582

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubaa/PCCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubaa/PCCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3013	100.0	1092	15	US-10-378-674-8
2	2810.5	93.3	892	12	US-10-282-122A-70481
3	2810.5	93.3	892	16	US-10-661-809-21
4	2810.5	93.3	930	12	US-10-615-383-10
5	2810.5	93.3	930	16	US-10-689-082-10
6	2810.5	93.3	930	16	US-10-690-184-10
7	2670	88.6	549	15	US-10-378-674-9
8	2510.5	83.3	560	15	US-10-378-674-2
9	1672	55.5	343	15	US-10-378-674-4
10	1562	51.8	316	15	US-10-378-674-6
11	1137.5	37.1	1141	12	US-10-282-122A-70251
12	1116.5	37.1	670	12	US-10-282-122A-70444
13	576	19.1	1349	9	US-09-815-242-5898
14	576	19.1	1349	9	US-09-815-242-13137
15	559	18.6	1385	12	US-10-282-122A-44324

16	499.5	16.6	932	9	US-09-815-242-5578	Sequence 5578, Ap
17	499.5	16.6	932	12	US-09-815-242-12438	Sequence 12438, A
18	499	16.6	1742	12	US-10-615-383-4	Sequence 4, Appli
19	499	16.6	1742	16	US-10-689-082-4	Sequence 4, Appli
20	499	16.6	1742	16	US-10-690-184-4	Sequence 4, Appli
21	493	16.4	1633	12	US-10-282-122A-70437	Sequence 70437, A
22	489	16.2	953	12	US-10-282-122A-44457	Sequence 44457, A
23	486.5	16.1	1920	12	US-10-282-122A-71413	Sequence 71413, A
24	455	15.1	841	9	US-09-815-242-5779	Sequence 5779, Ap
25	455	15.1	841	9	US-09-815-242-12751	Sequence 12751, A
26	454	15.1	936	8	US-08-781-986A-5249	Sequence 5249, Ap
27	454	15.1	936	12	US-10-329-624-5249	Sequence 5249, Ap
28	452	15.0	935	12	US-10-282-122A-44326	Sequence 44326, A
29	449	14.9	1021	9	US-09-815-242-5471	Sequence 5471, Ap
30	449	14.9	1021	9	US-09-815-242-12544	Sequence 12544, A
31	443.5	14.7	520	14	US-10-056-052-2	Sequence 2, Appli
32	397.5	13.2	877	12	US-10-282-122A-70428	Sequence 70428, A
33	377	12.5	496	14	US-10-311-879-29	Sequence 29, Appli
34	376	12.5	767	9	US-09-815-242-5899	Sequence 5899, Ap
35	376	12.5	767	9	US-09-815-242-13140	Sequence 13140, A
36	349	11.6	331	14	US-10-056-052-4	Sequence 4, Appli
37	349	11.6	345	9	US-09-813-820-7	Sequence 7, Appli
38	344.5	11.4	961	12	US-10-282-122A-43778	Sequence 43778, A
39	326	10.8	1038	12	US-10-282-122A-43827	Sequence 43827, A
40	319.5	10.6	978	9	US-09-815-242-5456	Sequence 5456, Ap
41	319.5	10.6	1001	9	US-09-815-242-12886	Sequence 12886, A
42	308.5	10.2	1018	9	US-09-815-242-5797	Sequence 5797, Ap
43	308.5	10.2	1018	9	US-09-815-242-12838	Sequence 12838, A
44	300.5	10.0	1027	8	US-08-781-986A-5254	Sequence 5254, Ap
45	300.5	10.0	1027	12	US-10-329-624-5254	Sequence 5254, Ap

ALIGNMENTS

RESULT 1

US-10-378-674-8  
; Sequence 8, Application US/10378674  
; Publication No. US20040006209A1  
; GENERAL INFORMATION:  
; APPLICANT: PARTI, Joseph M.  
; TITLE OF INVENTION: MONOCLONAL AND POLYCLONAL ANTIBODIES RECOGNIZING COAGULASE-NEGATI  
; FILE REFERENCE: P07556US01/BAS  
; CURRENT APPLICATION NUMBER: US/10/378,674  
; CURRENT FILING DATE: 2003-03-05  
; PRIOR APPLICATION NUMBER: 60/361,324  
; PRIOR FILING DATE: 2002-03-05  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 8  
; LENGTH: 1092  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-10-378-674-8

Query Match 100.0%; Score 3013; DB 15; Length 1092;  
Best Local Similarity 100.0%; Pred. No. 1.3e-176;  
Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	SSDEEKNDVNNQNSINTDDNNQIIKKEETNNYDGLKEKRSDETESTTNDENEATFLQK 60
Db	75	SSDEEKNDVNNQNSINTDDNNQIIKKEETNNYDGLKEKRSDETESTTNDENEATFLQK 134
Qy	61	TPQDNTHLTEEEVKESSESSVSSSSIDTAQPSHTTINREESVQTSNVDSDSHVSDPANS 120
Db	135	TPQDNTHLTEEEVKESSESSVSSSSIDTAQPSHTTINREESVQTSNVDSDSHVSDPANS 194
Qy	121	KIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNIDEXISNQDELNLNLPINEYENKAR 180
Db	195	KIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNIDEXISNQDELNLNLPINEYENKAR 254
Qy	181	PLSTTSAQPSIKRVTNQLAAEQGSNNVHLIKVTDQSITEGYDDSEGVIAKHAENLIYD 240

Db 255 PLSTTSAQPSIKRVTNQLAAEQGSNVNHLIKVTDQSITGEYDDSEGVKAHDAENLIYD 314  
Qy 241 VTFEVDKVKSGDVTWVIDKNTVPSDLTDSFTPIKIDNSGEIITATGTYNKNKQIITYT 300  
Db 315 VTFEVDKVKSGDVTWVIDKNTVPSDLTDSFTPIKIDNSGEIITATGTYNKNKQIITYT 374  
Qy 301 FTDYVDKYENIKAHKLKTSYIDSKVPNNNTKLDVEYKTALSSVNNKTITVEYQRPENRT 360  
Db 375 FTDYVDKYENIKAHKLKTSYIDSKVPNNNTKLDVEYKTALSSVNNKTITVEYQRPENRT 434  
Qy 361 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVIISGNGDEGSIIDSTIIKVKYVG 420  
Db 435 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVIISGNGDEGSIIDSTIIKVKYVG 494  
Qy 421 DNQNLPSNRIDYSEYEDVTNDDYAQLGNNDNVNINFGNIDSPVIIKVIKYPNKDDY 480  
Db 495 DNQNLPSNRIDYSEYEDVTNDDYAQLGNNDNVNINFGNIDSPVIIKVIKYPNKDDY 554  
Qy 481 TTIQQTVMQTTINEYTGFEFTASVDNTIAFSTSSGQGGDLPEKTYKIGDYVWEDVDK 540  
Db 555 TTIQQTVMQTTINEYTGFEFTASVDNTIAFSTSSGQGGDLPEKTYKIGDYVWEDVDK 614  
Qy 541 DGIQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDEGKYQFDG 582  
Db 615 DGIQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDEGKYQFDG 656

## RESULT 2

US-10-282-122A-70481  
; Sequence 70481, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 70481

; LENGTH: 892  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-10-282-122A-70481  
Query Match 93.3%; Score 2810.5; DB 12; Length 892;  
Best Local Similarity 93.6%; Pred. No. 2.8e-164;  
Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;  
Qy 1 SSDEKNDVNNQSIINTDDNNQIIKKEETNNYDGIKRSERDSTESTNVDHEATFLOK 60  
Db 35 SSNEEKNDVNNQSIINTDDNNQ-IKKEETNSDAIENRSKDTQSTNVDHEATFLOK 93  
Qy 61 TPQDNTLHTEEEVKESSESSVSSSIDTAQOPGHTTINREESVQTSNDVSDSHVSDFANS 120  
Db 94 TPQDNTLKEVEVKESSESSVSSSMDTAQOPGHTTINSEASQTSNDEENSRVSDFANS 153  
Qy 121 KIKESNTESKENTIEQPNKVKEDSTSPSGVTNIDEKISNQDELLNLPINEYENKAR 180  
Db 154 KIIESNTESKENTIEQPNKVEDSITSPSSVYKXIDEKISNQDELLNLPINEYENKVR 213  
Qy 181 PLSTTSAQPSIKRVTNQLAAEQGSNVNHLIKVTDQSITGEYDDSEGVKAHDAENLIYD 240  
Db 214 PLSTTSAQPSIKRVTNQLAAEQGSNVNHLIKVTDQSITGEYDDSGIITKAHDAENLIYD 273  
Qy 241 VTFEVDKVKSGDVTWVIDKNTVPSDLTDSFTPIKIDNSGEIITATGTYNKNKQIITYT 300  
Db 274 VTFEVDKVKSGDVTWVIDKNTVPSDLTDSFALPKIDNSGEIITATGTYNKNKQIITYT 333  
Qy 301 FTDYVDKYENIKAHKLKTSYIDSKVPNNNTKLDVEYKTALSSVNNKTITVEYQRPENRT 360  
Db 334 FTDYVDKYENIKAHKLKTSYIDSKVPNNNTKLDVEYKTALSSVNNKTITVEYQRPENRT 393  
Qy 361 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVIISGNGDEGSIIDSTIIKVKYVG 420  
Db 394 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVIISGNGDEGSIIDSTIIKVKYVG 453  
Qy 421 DNQNLPSNRIDYSEYEDVTNDDYAQLGNNDNVNINFGNIDSPYIIKVISKYDPNKDDY 480  
Db 454 DNQNLPSNRIDYSEYEDVTNDDYAQLGNNDNVNINFGNIDSPYIIKVISKYDPNKDDY 513  
Qy 481 TTIQQTVMQTTINEYTGFEFTASVDNTIAFSTSSGQGGDLPEKTYKIGDYVWEDVDK 540  
Db 514 TTIQQTVMQTTINEYTGFEFTASVDNTIAFSTSSGQGGDLPEKTYKIGDYVWEDVDK 573  
Qy 541 DGIQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDEGKYQFDG 582  
Db 574 DGIQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDEGKYQFDG 615

## RESULT 3

US-10-661-809-21  
; Sequence 21, Application US/10661809  
; Publication No. US20040101919A1  
; GENERAL INFORMATION:  
; APPLICANT: HOOK, Magnus  
; TITLE OF INVENTION: BIOINFORMATIC METHOD FOR IDENTIFYING SURFACE-ANCHORED PROTEINS FR  
; FILE REFERENCE: POSITIVE BACTERIA  
; FILE REFERENCE: P07741US01/BAS  
; CURRENT APPLICATION NUMBER: US/10/661,809  
; CURRENT FILING DATE: 2003-09-15  
; PRIOR APPLICATION NUMBER: 60/410303  
; PRIOR FILING DATE: 2002-09-13  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 21  
; LENGTH: 892  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-10-661-809-21  
Query Match 93.3%; Score 2810.5; DB 16; Length 892;  
Best Local Similarity 93.6%; Pred. No. 2.8e-164;



Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;

Qy 1 SSDEKNDVNNQSINTDDNQ-IKKEETNSDAIENRSKIDTQSTTNVDENEATFLOK 60  
Db 35 SSNEEKNDVNNQSINTDDNQ-IKKEETNSDAIENRSKIDTQSTTNVDENEATFLOK 93

Qy 61 TPQDNTLHTEEVKESSESSVSSSSIDTAQOPSHHTINREESVQTSNVDSDSHVDFPANS 120  
Db 94 TPQDNTQLKEEVKESSESSVSSSSMDTAQOPSHHTINSEASIQTSNDEENSRVDFPANS 153

Qy 121 KIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDDKISNODELNLPINEYENKAR 180  
Db 154 KIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDDKISNODELNLPINEYENKAR 213

Qy 181 PLSTTSAQPSIKRVTVNQLAAEQSGSNVNHILIKVTDQSIETGYDDSEGVIIKAHAENLIYD 240  
Db 214 PLSTTSAQPSIKRVTVNQLAAEQSGSNVNHILIKVTDQSIETGYDDSEGVIIKAHAENLIYD 273

Qy 241 VTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFTPIKIDNSGEIATGTVDNKNKQIYTT 300  
Db 274 VTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFTPIKIDNSGEIATGTVDNKNKQIYTT 333

Qy 301 FTDVVDKYENIKAKHLKTSYIDKSKVPNNNTKLDVEYKKTALSSVKNKTIITVEYQRPENRT 360  
Db 334 FTDVVDKYENIKAKHLKTSYIDKSKVPNNNTKLDVEYKKTALSSVKNKTIITVEYQRPENRT 393

Qy 361 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVINSNGDEGSIIDDDSTIIKVKYKVG 420  
Db 394 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVINSNGDEGSIIDDDSTIIKVKYKVG 453

Qy 421 DNQNLPSNRIDYSEYEDVTNDYDIAQLGNNDVNFNGNIDSPYIIKVISKYPNKDDY 480  
Db 454 DNQNLPSNRIDYSEYEDVTNDYDIAQLGNNDVNFNGNIDSPYIIKVISKYPNKDDY 513

Qy 481 TTIQQTVMQTTINEYTGERTASYDNTIAPSTSSGQGGDLPEKTYKIGDYVWEDVDK 540  
Db 514 TTIQQTVMQTTINEYTGERTASYDNTIAPSTSSGQGGDLPEKTYKIGDYVWEDVDK 573

Qy 541 DGIQNTNDNEKPLSNVLTLTYTPDGTSKSVRTDGDGKYQFDG 582  
Db 574 DGIQNTNDNEKPLSNVLTLTYTPDGTSKSVRTDGEKGYQFDG 615

RESULT 4  
US-10-615-383-10  
; Sequence 10, Application US/10615383  
; Publication No. US20040038327A1  
; GENERAL INFORMATION:  
; APPLICANT: FOSTER, Timothy  
; TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLOCOCCUS  
; FILE REFERENCE: P06335US03/BAS  
; CURRENT APPLICATION NUMBER: US/10/615,383  
; CURRENT FILING DATE: 2003-07-09  
; PRIOR APPLICATION NUMBER: 09/386,962  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: 60/098,443  
; PRIOR FILING DATE: 1998-08-31  
; PRIOR APPLICATION NUMBER: 60/117,119  
; PRIOR FILING DATE: 1999-01-25  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 10  
; LENGTH: 930  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-10-615-383-10

Query Match 93.3%; Score 2810.5; DB 12; Length 930;  
Best Local Similarity 93.6%; Pred. No. 3e-164;  
Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;

Qy 1 SSDEKNDVNNQSINTDDNQ-IKKEETNNYDGIKESRSDRTSTTNVDENEATFLOK 60

Db 73 SSNEEKNDVNNQSINTDDNQ-IKKEETNSDAIENRSKIDTQSTTNVDENEATFLOK 131

Qy 61 TPQDNTLHTEEVKESSESSVSSSSIDTAQOPSHHTINREESVQTSNVDSDSHVDFPANS 120  
Db 132 TPQDNTQLKEEVKESSESSVSSSSMDTAQOPSHHTINSEASIQTSNDEENSRVDFPANS 191

Qy 121 KIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDDKISNODELNLPINEYENKAR 180  
Db 192 KIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDDKISNODELNLPINEYENKAR 251

Qy 181 PLSTTSAQPSIKRVTVNQLAAEQSGSNVNHILIKVTDQSIETGYDDSEGVIIKAHAENLIYD 240  
Db 252 PLSTTSAQPSIKRVTVNQLAAEQSGSNVNHILIKVTDQSIETGYDDSEGVIIKAHAENLIYD 311

Qy 241 VTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFTPIKIDNSGEIATGTVDNKNKQIYTT 300  
Db 312 VTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFTPIKIDNSGEIATGTVDNKNKQIYTT 371

Qy 301 FTDVVDKYENIKAKHLKTSYIDKSKVPNNNTKLDVEYKKTALSSVKNKTIITVEYQRPENRT 360  
Db 372 FTDVVDKYENIKAKHLKTSYIDKSKVPNNNTKLDVEYKKTALSSVKNKTIITVEYQRPENRT 431

Qy 361 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVINSNGDEGSIIDDDSTIIKVKYKVG 420  
Db 432 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVINSNGDEGSIIDDDSTIIKVKYKVG 491

Qy 421 DNQNLPSNRIDYSEYEDVTNDYDIAQLGNNDVNFNGNIDSPYIIKVISKYPNKDDY 480  
Db 492 DNQNLPSNRIDYSEYEDVTNDYDIAQLGNNDVNFNGNIDSPYIIKVISKYPNKDDY 551

Qy 481 TTIQQTVMQTTINEYTGERTASYDNTIAPSTSSGQGGDLPEKTYKIGDYVWEDVDK 540  
Db 552 TTIQQTVMQTTINEYTGERTASYDNTIAPSTSSGQGGDLPEKTYKIGDYVWEDVDK 611

Qy 541 DGIQNTNDNEKPLSNVLTLTYTPDGTSKSVRTDGDGKYQFDG 582  
Db 612 DGIQNTNDNEKPLSNVLTLTYTPDGTSKSVRTDGEKGYQFDG 653

RESULT 5  
US-10-689-082-10  
; Sequence 10, Application US/10689082  
; Publication No. US20040142348A1  
; GENERAL INFORMATION:  
; APPLICANT: FOSTER, Timothy  
; TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLOCOCCUS  
; FILE REFERENCE: P06335US04/BAS  
; CURRENT APPLICATION NUMBER: US/10/689,082  
; CURRENT FILING DATE: 2003-10-21  
; PRIOR APPLICATION NUMBER: 09/386,962  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: 60/098,443  
; PRIOR FILING DATE: 1998-08-31  
; PRIOR APPLICATION NUMBER: 60/117,119  
; PRIOR FILING DATE: 1999-01-25  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 10  
; LENGTH: 930  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-10-689-082-10

Query Match 93.3%; Score 2810.5; DB 16; Length 930;  
Best Local Similarity 93.6%; Pred. No. 3e-164;  
Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;

Qy 1 SSDEKNDVNNQSINTDDNQ-IKKEETNNYDGIKESRSDRTSTTNVDENEATFLOK 60

Db 73 SSNEEKNDVNNQSINTDDNQ-IKKEETNSDAIENRSKIDTQSTTNVDENEATFLOK 131

Qy 61 TPQDNTLHTEEVKESSESSVSSSSIDTAQOPSHHTINREESVQTSNVDSDSHVDFPANS 120

Db 132 TPQDNTQLKEVYVKEPSSVSSNSMDTAQPSHTTINSEASIQTSDEENSRVDFPANS 191  
QY 121 KIKESNTESKGEENTIEQPNKVKEDSTTSOPSGYTNIDEKISNODELLNLPINEYENKAR 180  
Db 192 KIKESNTESKGEENTIEQPNKVKEDSTTSOPSGYTNIDEKISNODELLNLPINEYENKVR 251  
QY 181 PLSTTSAQPSIKRVTNQLAAEQGSNVNHLIKVTDQSITEGYDDSEGVIKAHDAENLIYD 240  
Db 252 PLSTTSAQPSKRVTVNQLAAEQGSNVNHLIKVTDQSITEGYDDSDGIIKAHDAENLIYD 311  
QY 241 VTFEVDKVKSGDVTVDIDKNTVPSDLTDSFTIPKIKDNGSGEIIATGTVDNKNKQIITYT 300  
Db 312 VTFEVDKVKSGDVTVDIDKNTVPSDLTDSFPAIPKIKDNGSGEIIATGTVDNKNKQIITYT 371  
QY 301 FTDYVDKYENIKAHKLKLTYSIDKSKVPNNNTKLDVEYKTALSSVNNKTIITVEYQRPENRT 360  
Db 372 FTDYVDKYENIKAHKLKLTYSIDKSKVPNNNTKLDVEYKTALSSVNNKTIITVEYQRPENRT 431  
QY 361 ANLQSMFTNIDTKNHTVQETIYINPLRYSAKETNNVNSGNGDEGSTIIDSTIIKVKYKVG 420  
Db 432 ANLQSMFTNIDTKNHTVQETIYINPLRYSAKETNNVNSGNGDEGSTIIDSTIIKVKYKVG 491  
QY 421 DNQNLPSNRIYDYSEYEDVTDNDYAQLGNNDVNFNGNIDSPYIIKVISKYDPNKDDY 480  
Db 492 DNQNLPSNRIYDYSEYEDVTDNDYAQLGNNDVNFNGNIDSPYIIKVISKYDPNKDDY 551  
QY 481 TTIOQTVMQTTINEYTGFEFTASVDNTIAFSTSGQGQGLPPEKTYKIGDYVWEDVDK 540  
Db 552 TTIOQTVMQTTINEYTGFEFTASVDNTIAFSTSGQGQGLPPEKTYKIGDYVWEDVDK 611  
QY 541 DGIQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDEGKYQFDG 582  
Db 612 DGIQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDEGKYQFDG 653

## RESULT 6

US-10-690-184-10

; Sequence 10, Application US/10690184

; Publication No. US20040141997A1

; GENERAL INFORMATION:

; APPLICANT: FOSTER, Timothy

; TITLE OF INVENTION: METHODS FOR TREATING OR PREVENTING INFECTIONS FROM COAGULASE-

; FILE REFERENCE: P0633US05/BAS

; CURRENT APPLICATION NUMBER: US/10/690,184

; CURRENT FILING DATE: 2003-10-21

; PRIOR APPLICATION NUMBER: 09/386,962

; PRIOR FILING DATE: 1999-08-31

; PRIOR APPLICATION NUMBER: 60/098,443

; PRIOR FILING DATE: 1998-08-31

; PRIOR APPLICATION NUMBER: 60/117,119

; PRIOR FILING DATE: 1999-01-25

; NUMBER OF SEQ ID NOS: 39

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 10

; LENGTH: 930

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

US-10-690-184-10

## Query Match

Best Local Similarity 93.3%; Score 2810.5; DB 16; Length 930;

Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;

QY 1 SSDEEKNDVINNNQSIINTDNNQIIKKEETNNYDGIKESRDETRTESTTNVDENEATFLOK 60  
Db 73 SSNEEKNDVINNNQSIINTDNNQIIKKEETNSDAIENRSKIDITQSTTNVDENEATFLOK 131  
QY 61 TPQDNTQLKEVYVKEPSSVSSNSMDTAQPSHTTINREESVQTSNDVEDSHVDFPANS 120  
Db 132 TPQDNTQLKEVYVKEPSSVSSNSMDTAQPSHTTINSEASIQTSDEENSRVDFPANS 191  
QY 121 KIKESNTESKGEENTIEQPNKVKEDSTTSOPSGYTNIDEKISNODELLNLPINEYENKAR 180

Db 192 KIKESNTESKGEENTIEQPNKVKEDSTTSOPSGYTNIDEKISNODELLNLPINEYENKVR 251  
QY 181 PLSTTSAQPSIKRVTNQLAAEQGSNVNHLIKVTDQSITEGYDDSEGVIKAHDAENLIYD 240  
Db 252 PLSTTSAQPSKRVTVNQLAAEQGSNVNHLIKVTDQSITEGYDDSDGIIKAHDAENLIYD 311  
QY 241 VTFEVDKVKSGDVTVDIDKNTVPSDLTDSFTIPKIKDNGSGEIIATGTVDNKNKQIITYT 300  
Db 312 VTFEVDKVKSGDVTVDIDKNTVPSDLTDSFPAIPKIKDNGSGEIIATGTVDNKNKQIITYT 371  
QY 301 FTDYVDKYENIKAHKLKLTYSIDKSKVPNNNTKLDVEYKTALSSVNNKTIITVEYQRPENRT 360  
Db 372 FTDYVDKYENIKAHKLKLTYSIDKSKVPNNNTKLDVEYKTALSSVNNKTIITVEYQRPENRT 431  
QY 361 ANLQSMFTNIDTKNHTVQETIYINPLRYSAKETNNVNSGNGDEGSTIIDSTIIKVKYKVG 420  
Db 432 ANLQSMFTNIDTKNHTVQETIYINPLRYSAKETNNVNSGNGDEGSTIIDSTIIKVKYKVG 491  
QY 421 DNQNLPSNRIYDYSEYEDVTDNDYAQLGNNDVNFNGNIDSPYIIKVISKYDPNKDDY 480  
Db 492 DNQNLPSNRIYDYSEYEDVTDNDYAQLGNNDVNFNGNIDSPYIIKVISKYDPNKDDY 551  
QY 481 TTIOQTVMQTTINEYTGFEFTASVDNTIAFSTSGQGQGLPPEKTYKIGDYVWEDVDK 540  
Db 552 TTIOQTVMQTTINEYTGFEFTASVDNTIAFSTSGQGQGLPPEKTYKIGDYVWEDVDK 611  
QY 541 DGIQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDEGKYQFDG 582  
Db 612 DGIQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDEGKYQFDG 653

## RESULT 7

US-10-378-674-9

; Sequence 9, Application US/10378674

; Publication No. US20040006209A1

; GENERAL INFORMATION:

; APPLICANT: PATTI, Joseph M.

; TITLE OF INVENTION: MONOCLONAL AND POLYCLONAL ANTIBODIES RECOGNIZING COAGULASE-NEGATIV

; FILE REFERENCE: P07556US01/BAS

; CURRENT APPLICATION NUMBER: US/10/378,674

; CURRENT FILING DATE: 2003-03-05

; PRIOR APPLICATION NUMBER: 60/361,324

; PRIOR FILING DATE: 2002-03-05

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 9

; LENGTH: 549

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

US-10-378-674-9

Query Match 88.6%; Score 2670; DB 15; Length 549;

Best Local Similarity 98.7%; Pred. No. 6.4e-156;

Matches 519; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSDEEKNDVINNNQSIINTDNNQIIKKEETNNYDGIKESRDETRTESTTNVDENEATFLOK 60  
Db 24 SSDEEKNDVINNNQSIINTDNNQIIKKEETNNNDGIEKSEDRTESTTNVDENEATFLOK 83  
QY 61 TPQDNTQLKEVYVKEPSSVSSNSMDTAQPSHTTINREESVQTSNDVEDSHVDFPANS 120  
Db 84 SPQDNTQLKEVYVKEPSSVSSNSMDTAQPSHTTINREESVQTSNDVEDSHVDFPANS 143  
QY 121 KIKESNTESKGEENTIEQPNKVKEDSTTSOPSGYTNIDEKISNODELLNLPINEYENKAR 180  
Db 144 KIKESNTESKGEENTIEQPNKVKEDSTTSOPSGYTNIDEKISNODELLNLPINEYENKAR 203  
QY 181 PLSTTSAQPSIKRVTNQLAAEQGSNVNHLIKVTDQSITEGYDDSEGVIKAHDAENLIYD 240  
Db 204 PLSTTSAQPSIKRVTNQLAAEQGSNVNHLIKVTDQSITEGYDDSEGVIKAHDAENLIYD 263

Qy 241 VTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFTPIPKIKNSGEIIATGTVDNKNKQIYTT 300  
Db 264 VTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFTPIPKIKNSGEIIATGTVDNKNKQIYTT 323  
Qy 301 FTDYVDKYENIKAKHLKLTSTYIDKSKVNNNTKLDVEYKTKALSSVKNKTIITVEYQRPENRT 360  
Db 324 FTDYVDKYENIKAKHLKLTSTYIDKSKVNNNTKLDVEYKTKALSSVKNKTIITVEYQRPENRT 383  
Qy 361 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVTNINSGNGDEGSTIIDSTIIKVKYKVG 420  
Db 384 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVTNINSGNGDEGSTIIDSTIIKVKYKVG 443  
Qy 421 DNQNLPSNRIYDYSEYEDVTDNDYAQLGNNDVNINFGNIDSPYIIKVIISKYDPNKDDY 480  
Db 444 DNQNLPSNRIYDYSEYEDVTDNDYAQLGNNDVNINFGNIDSPYIIKVIISKYDPNKDDY 503  
Qy 481 TTIQQTVMQTINNEYTGERTASYDNTIAFSTSSGQGGDLPEK 526  
Db 504 TTIQQTVMQTINNEYTGERTASYDNTIAFSTSSGQGGDLPEK 549

RESULT 8  
US-10-378-674-2  
; Sequence 2, Application US/10378674  
; Publication No. US20040006209A1  
; GENERAL INFORMATION:  
; APPLICANT: PATTI, Joseph M.  
; TITLE OF INVENTION: MONOCLONAL AND POLYCLONAL ANTIBODIES RECOGNIZING COAGULASE-NEGAT  
; FILE REFERENCE: P07556US01/BAS  
; CURRENT APPLICATION NUMBER: US/10/378, 674  
; CURRENT FILING DATE: 2003-03-05  
; PRIOR APPLICATION NUMBER: 60/361,324  
; PRIOR FILING DATE: 2002-03-05  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 560  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-10-378-674-2

Query Match 83.3%; Score 2510.5; DB 15; Length 560;  
Best Local Similarity 93.2%; Pred. No. 4.1e-146;  
Matches 490; Conservative 14; Mismatches 21; Indels 1; Gaps 1;

Qy 1 SSDEKNDVINNQSINTDDNNQIIKKEETNNYDIEKRSERDTESTTNDENEATFLOK 60  
Db 36 SSNEEKNDVINNQSINTDDNQ- IKKEETNSDAIENRSKDIQTOSTTNDENEATFLOK 94  
Qy 61 TPQDNTHLTBEEVKSSSVSSSSIDTAQPSHTTINREESVQTSNDVSDSHVDFANS 120  
Db 95 TPQDNTQLKEEVKPSVSSVSSSSMDTAQPSHTTINSEASIQTSNDSENSRVSDFANS 154  
Qy 121 KIIESNTESKEENTIIQPNKVKEDSTTSQPSGTYNIDEKISNODELLNLPINEYENKAR 180  
Db 155 KIIESNTESKEENTIIQPNKVKEDSTTSQPSGYNIDEKISNODELLNLPINEYENKVR 214  
Qy 181 PLSTTSAPQSKRTVTNQLAAEQSNVNHIIKVTQDSITTEGYDDSEGVKHAHAENLIYD 240  
Db 215 PLSTTSAPQSKRTVTNQLAAEQSNVNHIIKVTQDSITTEGYDDSDGIIKAHAENLIYD 274  
Qy 241 VTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFTPIPKIKNSGEIIATGTVDNKNKQIYTT 300  
Db 275 VTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFTPIPKIKNSGEIIATGTVDNKNKQIYTT 334  
Qy 301 FTDYVDKYENIKAKHLKLTSTYIDKSKVNNNTKLDVEYKTKALSSVKNKTIITVEYQRPENRT 360  
Db 335 FTDYVDKYENIKAKHLKLTSTYIDKSKVNNNTKLDVEYKTKALSSVKNKTIITVEYQRPENRT 394  
Qy 361 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVTNINSGNGDEGSTIIDSTIIKVKYKVG 420  
Db 395 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVTNINSGNGDEGSTIIDSTIIKVKYKVG 454

Qy 421 DNQNLPSNRIYDYSEYEDVTDNDYAQLGNNDVNINFGNIDSPYIIKVIISKYDPNKDDY 480  
Db 455 DNQNLPSNRIYDYSEYEDVTDNDYAQLGNNDVNINFGNIDSPYIIKVIISKYDPNKDDY 514  
Qy 481 TTIQQTVMQTINNEYTGERTASYDNTIAFSTSSGQGGDLPEK 526  
Db 515 TTIQQTVMQTINNEYTGERTASYDNTIAFSTSSGQGGDLPEK 560

RESULT 9  
US-10-378-674-4  
; Sequence 4, Application US/10378674  
; Publication No. US20040006209A1  
; GENERAL INFORMATION:  
; APPLICANT: PATTI, Joseph M.  
; TITLE OF INVENTION: MONOCLONAL AND POLYCLONAL ANTIBODIES RECOGNIZING COAGULASE-NEGAT  
; FILE REFERENCE: P07556US01/BAS  
; CURRENT APPLICATION NUMBER: US/10/378, 674  
; CURRENT FILING DATE: 2003-03-05  
; PRIOR APPLICATION NUMBER: 60/361,324  
; PRIOR FILING DATE: 2002-03-05  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 343  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-10-378-674-4

Query Match 55.5%; Score 1672; DB 15; Length 343;  
Best Local Similarity 97.9%; Pred. No. 7.1e-95;  
Matches 319; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 201 AEOGSNVNHLIKVTDQISITEGYDDSEGVKHAHAENLIYDVFEDDKVSGDGMTVTDID 260  
Db 18 SEQGSNVNHLIKVTDQISITEGYDDSGIILKAHAENLIYDVFEDDKVSGDGMTVTDID 77  
Qy 261 KNTVPSDLTDSFTPIPKIKNSGEIIATGTVDNKNKQIYTTFTDYVDKYENIKAKHLKLTSTY 320  
Db 78 KNTVPSDLTDSFPAIPKIKNSGEIIATGTVDNKNKQIYTTFTDYVDKYENIKAKHLKLTSTY 137  
Qy 321 IDKSKVPNNNTKLDVEYKTKALSSVKNKTIITVEYQRPENRTANLQSMFTNIDTKNHTVEQT 380  
Db 138 IDKSKVPNNNTKLDVEYKTKALSSVKNKTIITVEYQRPENRTANLQSMFTNIDTKNHTVEQT 197  
Qy 381 IYINPLRYSAKETNVTNINSGNGDEGSTIIDSTIIKVKYKVDNQNLPDSNRIYDYSEYEDV 440  
Db 198 IYINPLRYSAKETNVTNINSGNGDEGSTIIDSTIIKVKYKVDNQNLPDSNRIYDYSEYEDV 257  
Qy 441 TNDYDQAQLGNNDVNINFGNIDSPYIIKVIISKYDPNKDDYTTTIQQTVMQTINNEYTGTF 500  
Db 258 TNDYDQAQLGNNDVNINFGNIDSPYIIKVIISKYDPNKDDYTTTIQQTVMQTINNEYTGTF 317  
Qy 501 RTASYDNTIAFSTSSGQGGDLPEK 526  
Db 318 RTASYDNTIAFSTSSGQGGDLPEK 343

RESULT 10  
US-10-378-674-6  
; Sequence 6, Application US/10378674  
; Publication No. US20040006209A1  
; GENERAL INFORMATION:  
; APPLICANT: PATTI, Joseph M.  
; TITLE OF INVENTION: MONOCLONAL AND POLYCLONAL ANTIBODIES RECOGNIZING COAGULASE-NEGAT  
; FILE REFERENCE: P07556US01/BAS  
; CURRENT APPLICATION NUMBER: US/10/378, 674  
; CURRENT FILING DATE: 2003-03-05  
; PRIOR APPLICATION NUMBER: 60/361,324  
; PRIOR FILING DATE: 2002-03-05

```
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-378-674-6

Query Match          51.8%; Score 1562; DB 15; Length 316;
Best Local Similarity 97.7%; Pred. No. 3.7e-88;
Matches 298; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 201 AEOGNNVHLIKVTDQSTIEGVDSGVKHAENLIYDVFVDDKVKSGDWTMTVDID 260
Db 12 SEQGNNVHLIKVTDQSTIEGVDSGVKHAENLIYDVFVDDKVKSGDWTMTVDID 71
QY 261 KNTVPSDLTDSFTTPKIKDNGSEIATGTYNKKNQIYTFYDVKYENIKAHKLKTSY 320
Db 72 KNTVPSDLTDSFAPKIKDNGSEIATGTYNKKNQIYTFYDVKYENIKAHKLKTSY 131
QY 321 IDKSKVPNNNTKLDVEYKTLSSVNTIIVYQRPNNRNTANLQSMFTNIDTKNHTVEQT 380
Db 132 IDKSKVPNNNTKLDVEYKTLSSVNTIIVYQRPNNRNTANLQSMFTNIDTKNHTVEQT 191
QY 381 IYINPLRYSAKETNNVNSGDEGSDIIDDSTIIKVKYKVDNQNLPSNRIDYSEYEDV 440
Db 192 IYINPLRYSAKETNNVNSGDEGSDIIDDSTIIKVKYKVDNQNLPSNRIDYSEYEDV 251
QY 441 TNDYAOAGNNNDVNIINFGNIDSPYIIKVISKYPNKKDDYTTIOQVTMTTINEYTGFE 500
Db 252 TNDYAOAGNNNDVNIINFGNIDSPYIIKVISKYPNKKDDYTTIOQVTMTTINEYTGFE 311
QY 501 RTASY 505
Db 312 RTASY 316

RESULT 11
US-10-282-122A-70251
; Sequence 70251, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EUITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70251
; LENGTH: 1141
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-282-122A-70251

Query Match          37.8%; Score 1137.5; DB 12; Length 1141;
Best Local Similarity 43.3%; Pred. No. 2.3e-61;
Matches 266; Conservative 97; Mismatches 193; Indels 59; Gaps 19;

QY 1 SSDEEKNDVNNQSGINTDDNNOIIEKEETNNYDGIKRESEDTSTNTVDENEATPLOK 60
Db 69 SDNKEVVSETEENSTTENDSTNFI--KKEINT--DSQPEAKEESTTSTQOQNNVTATTE 125
QY 61 TPQDNTHLTFEEFVKES----SSVESNSSIDTAQPSHTTINREESVQTSNVEDSHVSDF 117
Db 126 TKPQN--IEKENVKPSTDKTATEDTSVILEEKAPNYT--NNDVTTKPS----- 170
QY 118 ANSKIKESNT--ESGKEENTIEQPNKVKEDSTTSOPSGYTNIDKIS--NODELLNLP-- 171
Db 171 -TSEIQTKPTTPOESTNIENSQPTPSKVD---NQVDTATNPKPEPVNVSKELKNPEK 226
QY 172 -----INEYENKARPLST--TSAQP-----STKRYTVNQLAAEOGSNNVHLIKVTDQSI 218
Db 227 LKELVRNDNNTDRSTKPVATPTSPVAPKRLNAKMRFAVAQPAASVANNVNDLITVTKQTI 286
QY 219 TEGYDDSEGVKHAENLIYDVFVDDKVKSGDWTMTVDIDKNTVPSDLTDSFTTPKIK 278
Db 287 KVG-DGKDNVAAHDKGDISEYDTEFTIDNKVKKGDTMTINYNKVNTPSOLDTKNDPIDIT 345
QY 279 DNSGEIATGTYNKKNQIYTFYDVKYENIKAHKLKTSYDVKSPNNNTKLDVEYK 338
Db 346 DPSEVIAGTFPKATKQIYTFYDVKYEDIKARLTLYSIDKQAVP-NETSLMTFA 404
QY 339 TALSSVNTIIVYQRPNNRNTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVIS 398
Db 405 TAGKETSONVSDYQDPMVHGDSNIQSIFTKLDENKQITTEQIYVNPFLKKTATNTKVDIA 464
QY 399 GNG-----DEGSTIIDDSTIIKVKYKVDNQNLPSNRIDYSEYEDVND-DYAO 448
Db 465 GSQVDDYGNIKLNGSTIIDDQTEIKVKYKVNPNQQLPQSNRIYDFSQYEDVTSQFNNKKS 524
QY 449 GNNNDVNIINFGNIDSPYIIKVISKYPNKKDDYTTIOQVTMTTINEYTGFEFTASYDNT 508
Db 525 FSNVATLDFGDNINSAYIIKVKSKYTPSTDGELDIAQGTSMRTT-DKY-GYNNYAGYSNF 582
QY 509 IAPSTSSGGQGGDL-PPERTYKIGDVYVWEDVDKGGIQTNDNEKPLSNVLTLTYPDGTS 567
Db 583 IVTSNDTGGGDTGTVKPEEKLYKIGDVYVWEDVDKGGVQGTDSKEKPMANVLTLTYPDGTT 642
QY 568 KSVRTDEDCGYOPDG 582
Db 643 KSVRTDANGHYEFGG 657

RESULT 12
US-10-282-122A-70444
; Sequence 70444, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
```

```
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-09-06
; PRIOR FILING DATE: 2000-09-09
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2001-02-16
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70444
; LENGTH: 670
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-10-282-122A-70444

Query Match 37.1%; Score 1116.5; DB 12; Length 670;
Best Local Similarity 98.2%; Pred. No. 2.3e-60;
Matches 213; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

Qy 366 MFTNIDPKNHTVEQTIYINPLRYSAKETNVSISNGDEGSIIDDSIIKVKYKVDNQNL 425
Db 1 MFTNIDPKNHTVEQTIYINPLRYSAKETNVSISNGDEGSIIDDSIIKVKYKVDNQNL 60

Qy 426 PDSNRIYDYSEYEDVTNDYDQAQLGNNDVNIINFGNIDSPYIIKVKYKVDNQDYYTIIQ 485
Db 61 PDSNRIYDYSEYEDVTNDYDQAQLGNNDVNIINFGNIDSPYIIKVKYKVDNQDYYTIIQ 120

Qy 486 TVTMOTTINVTGFRFASDNTIAFSTSGOGGDLPPPEKTYKIGYVWEDVDKQIGON 545
Db 121 TVTMOTTINVTGFRFASDNTIAFSTSGOGGDLPPPEKTYKIGYVWEDVDKQIGON 177

Qy 546 TNDNEKPLSNVLTYPDGTGSKSVRTDEDKGYFDG 582
Db 178 TNDNEKPLSNVLTYPDGTGSKSVRTDEDKGYFDG 214

RESULT 13
US-09-815-242-5898
; Sequence 5898, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haeelbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5898
; LENGTH: 1349
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-09-815-242-5898

Query Match 19.1%; Score 576; DB 9; Length 1349;
Best Local Similarity 30.1%; Pred. No. 9.2e-27;
Matches 186; Conservative 112; Mismatches 228; Indels 92; Gaps 29;

Qy 1 SDEERKNDV---INNQSINTDNNQIIKKEETNNVDGIEKRSDETRTESTNVDNE-NEAT 56
Db 55 STNKLNEATTSASDNQSSDKVDQMLNQEDNTKNDKQEMVSSQGNETTSGNKLIERE 114

Qy 57 FLQKTPQDNTHLTTEEVEKSSSVESNSSIDTAQPSHTTINEESVQTSDDNVDHVSVD 116
Db 115 SVQSTTGKVKVSTAKSDQASPKSTNEDLNTKQ----TISNQEALQ-PDLQENKSVVN 168

Qy 117 FANSKIKESNTESKEBENTIEQPNKVKEDSTTSQPSGYTNI-DEKISNQDELL--NLPIN 173
Db 169 -----VQPTNEENKVD-----AKTESTT-----LNVKSDAISKNDLVDNNSNS 209

Qy 174 EYENKARPSTTSQAQP-----SIKRVTVNQLAABQGNVNNHLIKVTDQSITEGVDDSEGI 229
Db 210 NNENNADIILPKSTAPKRLNTRMRIAQVSPSTEAKNVNDLITSNTTLTVVDADKNNKIV 269

Qy 230 KAHDANLIYDVTPEYDDVKSGDTMTVDIDKNTV-----PSDLTDSFTPIPKIKO-NSG 282
Db 270 PAQDYLKSLKQIT--VDDKVKSGDYFTIKY-SDTVQVYGLNPEDIKN---IGDIKDPNG 323

Qy 283 EITATGYDNKKNQIITYFTFDYVDKYENIKAKHLKLTSYIDKSKVPNNNTKLDVEYKITALS 342
Db 324 ETIATAKHTANNLIITYFTFDYVDKFNQVGMINGYSYMDADTIP--VSKNDVEFNVTIG 381

Qy 343 SVNKTITVEYQRPN--ENRTANLQSMFTNIDTKNHT-----VEQTIYINPLRYSAKE 392
Db 382 NTTTTKTANIQYDPYVNEKNKSGSAFT--ETVSHVGNKENPGYKQTIYVNPSENSLTN 439

Qy 393 TYNVI-----SGNGDEGSTIIDDSIIKVKYKVDNQNLPSNRIYDYS--EVEDVNDYD 446
Db 440 AKLKVQAYHSSYPNNITGQINKVDYDIKIQVPGKGYTL---NKGVDVNTKELTDVTN-QYL 495

Qy 447 Q---LGNNDVNIINFGNIDSPYIIKVKYKVDNQDYYTIIQQTVMOTTINEYTGPERTA 503
Db 496 QKITYGDNNSAVIDFGNADSAVYVWNTKEQYTNSESPTLVQMTLSSTGN-----KSV 549

Qy 504 SYDNTIAFSTSGOGGDLPPPEKTYKIGYVWEDVDKQIGONTDNNEKPLSNVLTYP 563
Db 550 STGNALGFTNNQSGGAG---QEVYKIGYVWEDVDKQIGQEL--GEKGVGNVTVTY-FD 602

Qy 564 DGTSKSVR---TDEDGKY 578
```









GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:42:37 ; Search time 21.2079 Seconds  
(without alignments)  
2639.748 Million cell updates/sec

Title: US-10-806-288-13  
Perfect score: 3013  
Sequence: 1 SSDEKNDVINNNQSIINTDD.....PDGTSKSVRTDEBQKQFDG 582

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3013	100.0	1092	2 T30214	fibrinogen-binding
2	1144.5	38.0	1166	2 T28680	fibrinogen-binding
3	1137.5	37.8	1141	2 E89824	hypothetical prote
4	576	19.1	1315	2 T28679	fibrinogen-binding
5	559	18.6	1385	2 D89824	hypothetical prote
6	489	16.2	953	2 C89824	hypothetical prote
7	452	15.0	933	2 S41539	fibrinogen-binding
8	452	15.0	989	2 D89852	fibrinogen-binding
9	397.5	13.2	877	2 P90070	Clumping factor B
10	376	12.5	940	2 S19702	fibronectin-bindin
11	344.5	11.4	961	2 G90053	hypothetical prote
12	326	10.8	1038	2 H90053	hypothetical prote
13	300.5	10.0	1018	2 A32192	fibronectin-bindin
14	239	7.9	1039	2 T30856	protein F2 - Strept
15	225	7.5	1463	2 T30280	AAS surface protei
16	216.5	7.2	1959	2 A31085	hypothetical prote
17	216.5	7.2	3394	2 T18501	hypothetical prote
18	212.5	7.1	1125	2 E90598	membrane nucleas
19	211.5	7.0	2206	2 G71611	hypothetical prote
20	208.5	6.9	1072	2 A86827	hypothetical prote
21	208	6.9	1127	2 T28317	ORF MSV156 hypothe
22	208	6.9	3724	2 T18427	hypothetical prote
23	207.5	6.9	1711	2 T18429	hypothetical prote
24	207.5	6.9	2401	2 T28676	rhodopy protein -
25	207	6.9	769	2 P89870	serine proteinase
26	207	6.9	1939	2 T18372	repeat organellar
27	207	6.9	4550	2 T18440	hypothetical prote
28	206.5	6.9	821	2 S67087	hypothetical prote
29	206	6.8	1302	1 JC6009	surface-located me

ALIGNMENTS

RESULT 1

T30214  
fibrinogen-binding protein - Staphylococcus epidermidis  
C;Species: Staphylococcus epidermidis  
C;Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 11-May-2000  
C;Accession: T30214  
R;Nilsson, M.; Frykberg, L.; Flock, J.I.; Pei, L.; Lindberg, M.; Guss, B.  
Infect. Immun. 66, 2666-2673, 1998  
A;Title: A Fibrinogen-binding protein of Staphylococcus epidermidis.  
A;Reference number: Z20781; MUID:98261511; PMID:9596732  
A;Accession: T30214  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-1092 <NIL>  
A;Cross-references: EMBL:Y17116; NID:el296734; PID:el296735; PIDN:CAA76638.1

Query Match 100.0%; Score 3013; DB 2; Length 1092;  
Best Local Similarity 100.0%; Pred. No. 7.2e-123;  
Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	SSDEKNDVINNNQSIINTDDNNQIIKKEETNNYDGIKRSSEDRTESTTNVDENATFLQK	60
Db	75	SSDEKNDVINNNQSIINTDDNNQIIKKEETNNYDGIKRSSEDRTESTTNVDENATFLQK	134
Qy	61	TPQDNTHLTTEEYKESSESSNSSIDTAQPSHTTINREESVQTSNVEDSHVDFANS	120
Db	135	TPQDNTHLTTEEYKESSESSNSSIDTAQPSHTTINREESVQTSNVEDSHVDFANS	194
Qy	121	KIKESNTESGKEENTIEQPNKVEDSTTSQPSGYTNIDEKISNQDELINLPINYEYENKAR	180
Db	195	KIKESNTESGKEENTIEQPNKVEDSTTSQPSGYTNIDEKISNQDELINLPINYEYENKAR	254
Qy	181	PLSTSAQPSIKRVTVNQLAAEGSNVNHILIKVTDQSIITGYDDSEGVKAHDAENLIYD	240
Db	255	PLSTSAQPSIKRVTVNQLAAEGSNVNHILIKVTDQSIITGYDDSEGVKAHDAENLIYD	314
Qy	241	VTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFTIPKIDNSGEIATGTVDNKNKQIYTT	300
Db	315	VTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFTIPKIDNSGEIATGTVDNKNKQIYTT	374
Qy	301	FTDYVDKYENIKAHKLKTSYIDKSVPNNNTKLDVEYKLTALSSVNKTIITVEYQPNENRT	360
Db	375	FTDYVDKYENIKAHKLKTSYIDKSVPNNNTKLDVEYKLTALSSVNKTIITVEYQPNENRT	434
Qy	361	ANQSQMFTNIDTKNHTVEQIYINPLRYSAKETNNVNSGNGDEGSTIIDSTIIKVKVKG	420
Db	435	ANQSQMFTNIDTKNHTVEQIYINPLRYSAKETNNVNSGNGDEGSTIIDSTIIKVKVKG	494
Qy	421	DNQNLPSNRIYDYSEYEDVTNDDYAQLGNNDNVNINFGNIDSPYIIKVISKYPNKKDDY	480
Db	495	DNQNLPSNRIYDYSEYEDVTNDDYAQLGNNDNVNINFGNIDSPYIIKVISKYPNKKDDY	554

Imp1 protein - Myc  
hypothetical prote  
FntB protein [mpo  
asparagine-rich pr  
DNA helicase II BR  
rhodopy protein -  
1-phosphatidylinos  
hypothetical prote  
hypothetical prote  
hypothetical prote  
similar to late em  
hypothetical prote  
probable membrane  
internalin protein  
hypothetical prote





```

Query Match      15.0%; Score 452; DB 2; Length 989;
Best Local Similarity 25.9%; Pred. No. 2,3e-12;
Matches 152; Conservative 109; Mismatches 224; Indels 102; Gaps 27;

QY      1 SSDEKKNVNNQSIINTDDNNQLIKKEETNNYDGIKRSEDRTSTTNVDENAATFLQK 60
       |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :|

```

```
Db 51 SNESKNSDSSVSAPKTDITNV-----SDTKSSNTN---NGETSVAQ 91
Qy 61 TPQDNTHLTHEEVKSSSVSSNSIDTAQPSHTTINREESVQTSNDVEDSHVDFPANS 120
Db 92 NP-----AQETTSQSSNTATBETPVTEGATTTTNTQANTPATQS-SNTNAEELVN- 143
Qy 121 KIKESNTESGKEENTIQPNKVKEDSSTSPSGVNTNDEKISNODELLN-LPINEYENKA 179
Db 144 --QTSNETTNDITNV-----SSVNSPQNSNAENVSTTQDTSTEATPSN---NES 189
Qy 180 RPLSTISA-----QPSIKRVTVNQLAA-----EQGSNVNHLIKVTDQOSITEGVDD 224
Db 190 APQNTDASNKDVQAVNPSTPRRAFSLAADAAPAGTDITN--QLTDVKVT---ID 244
Qy 225 SEGVIKAHAENLIYDVTFFVDDKVKGDNTVTDIDKNTVPSDLTDSFTPIKIKDNGSEI 284
Db 245 SGTTVYHQAGVYKLVNGFSPNSAVKGTFTKIVPKELNGLVSTAKVPPIMAGD-QV 303
Qy 285 IATCTYONKQIYTFDVTYDVKYENIKAHKLTSYDKSKVPNNTKL-DVEYKTL-- 341
Db 304 LANGVIDS-DGNVITYTFDYDVKENVTANITMPAYID----PENVTKTGNVTLTTGIGT 358
Qy 342 SSVAKTITVEYQRPENRNTALQSMFTNIDTKNHTVEQTIYINPLRYSAKETNV----- 395
Db 359 NTASKTVLIDYEKGQFNLISIKGTIDQIDKNTNYRQTIYVNP-----SGDNVVLPAIT 413
Qy 396 -NISGNDGEGSTIIDSTIIKVKYVGNQNLPLDSNRIYDYSEYEDVTDNDYDVAQLGNNDV 454
Db 414 GNLPNTKSNALIDAKNTDIIKVRV-DWANDLSSYYVNPSPDFVTVNQVRISPPNANQY 472
Qy 455 NINF-----GNLDSPIIKVSKYDPKDDYTIQCTVMTQTTINEYTGEP--RTASVDNT 508
Db 473 KVEFPDQDDQITPYIVVNGHIDPASTG-----DLALRSTFYGDSNFWRSMSWDNE 526
Qy 509 IAFSTSSQGGG-DLP--PEKTYKIGDY--VVEDVDKD-GTQNTNDN 549
Db 527 VAFNNGSGGIDKPKVPVPEQDFRGEIEPIPEDSDPGSDSGSDS 573

RESULT 9
F90070
Clumping factor B [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: F90070
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc-
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: F90070
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-877 <KUP>
A:Cross-references: GB:BA000018; PID:gl3702588; PIDN:BA843728.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: clfB

Query Match 13.2%; Score 397.5; DB 2; Length 877;
Best Local Similarity 24.5%; Pred. No. 4.4e-10;
Matches 143; Conservative 81; Mismatches 207; Indels 153; Gaps 23;

Qy 2 SDEKNDVI---NNQSIINTDNNQI IKKETNNYDGIKESEDRTESTTNNVDNEATFL 58
Db 44 ASEQSDNTTQSKKNASADSEKNMI-----ETPQLTNTANDTSDISANTNSANVDSTTKP 99
Qy 59 QKTPODNTHLTHEEVKSSSVSSNSIDTAQPSHTTINREESVQTSNDVEDSHVSDFA 118
Db 100 MSTQTSNTTTE-----PASTNETPQTAIKNQ---ATAAKMQDQTVQEA 142
Qy 119 NSKIKESNTESGKEENTIEQPNKVKEDSTTSPSGVNTNDEKISNODELLNLPINEYENK 178
```

```
Db 143 NSQVDNKT---NDANSIATNSLKNST-----LDLP-----Q 173
Qy 179 ARPLSTTSAQ-----PSIKRVTVNQLAAEQ-----GSNVNHLIKVTDQOSITEGVDD 225
Db 174 SSQPTISNAQGTSPKSPVTRAVSLAABPVVAADAKGTNVN-----DKVTASNEKLE 227
Qy 226 EGVIKAHAENLIYDVTFFVDDKVKGDNTV-----DID-----KNTVPSDLTDS 271
Db 228 KTTFDENQSGNTFMAANFTVTDKVKSGDYFTAKLPDSLGTGNGDVDSNSNTMP----- 281
Qy 272 FTIPKIKDNGSEIATCTYONKQIYTFDVTYDVKYENIKAHKLTSYDKSKVPNNNT 331
Db 282 --IADIKSTNGDVVAKATYDILTKTYTFVFDVNNKENINGQFSLPLFTRAKAPKSGT 339
Qy 332 KLDVEYKALTALSSVNTKITVEYQRP-----NENRNTANLQSMFTNIDTKN--HTVEQTIYINP 385
Db 340 -YDANINIADEMFPNNKITVYSPFIAGIDKPGCANISSIIIGVDTASGQNTYKQTVFVNP 398
Qy 386 LRYSAKETNVNISGNDGEGSTIIDDS-----TIKVKYVGNQNLPLDS----- 428
Db 399 KORVLGNTWYIIRGYQDK-----IESSGKVSATDTKLRIFEVNDTSLSDSYADPNDSN 454
Qy 429 -----NRIYDYSYEDVTDNDYDVAQLGNNDVNINFGNIDSPYIIKISKYDPNKDD 479
Db 455 LKEVTDQFKNRIY-----YE-----HPNVASIKFGDITKYTVVLVEGHYDNTGKN 499
Qy 480 YTT--IQCTVMTQTTINEYTGEPRTASYDNTIAFSTSSQGGQGD 521
Db 500 LKTVQIQENVDPVT-----NRDYSIFGNNNENVVRYGGGSADGD 538

RESULT 10
S19702
fibronectin-binding protein B - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Oct-1999
C:Accession: S19702
R:Joensson, K.; Signaes, C.; Mueller, H.P.; Lindberg, M.
Eur. J. Biochem. 202, 1041-1048, 1991
A:Title: Two different genes encode fibronectin binding proteins in Staphylococcus aureus
A:Reference number: S19702; MUID:92111475; PMID:1837266
A:Accession: S19702
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-940 <JOE>
A:Cross-references: EMBL:X62992; NID:g49040; PIDN:CAA44726.1; PID:g581562
C:Keywords: fibronectin binding

Query Match 12.5%; Score 376; DB 2; Length 940;
Best Local Similarity 23.4%; Pred. No. 4.1e-09;
Matches 138; Conservative 110; Mismatches 217; Indels 124; Gaps 22;

Qy 35 GIEKRSERDRTSTNVNDNEATFLQKTPQDNTHLTHEEVKSSSVSSNSIDTAQ--- 90
Db 30 GQEKAAASEQNNTTVEES-----GSSATESKASSETQTNTNNVTIDTQSYSA 78
Qy 91 ----QPSHTT-INREESVQT--SDNVEDSHVSDFANSKIKESNTESGKEENTIEQPNKVK 143
Db 79 TSTEQPSQSTQVTTTEEAPKTQAPKVSrv-DLPSEKADKETTGTQVD--IAQPSNV 135
Qy 144 EDSTTSPSGVNTNDEKISNODELLNLPINEYENKAPLSTTSAQPSIKRVTVNQLAAEQ 203
Db 136 E-----IKPRKGRSTDVTAVEK 153
Qy 204 -----GSNVNHLIKVTDQSITEGVDDSEGVIKAHDAENLIYDVTFFVDDKVKSGDT 254
Db 154 EVVEETKATGTDVTKNVEGSEIGVGHKQDNTVNVNPNHAEVTLKVKFKPGEIGKAGDY 213
Qy 255 MTVDIDKNTVPSDLTDSFTPIKIKDNGSEIATCTYONKQIYTFDVTYDVKYENIKAH 314
Db 214 PFTLSNDVETHGISTLRKVPKISTDQVWATGEIIGERK-VRYTFKEYVQEKKDLTAE 272
```

```
QY 315 LKLTSDVSKVPNNNTKLDVEYKTALESVNKTIIVFY---QRPNNRNTANLQSMFTNID 371
DB 273 LSLNLFIDPTTQKNGO-NVEVGLGETTQSVKIFNIQILGVRDNGWVTAN-----GRID 326
QY 372 TKNHT---VEQTIYINPLRYSAKETNVAISNGSDGSTIIDSTIIKVKYKVDNQNLPS 428
DB 327 TLNKVDGKSFHAYMKPNQSL--SSVTVTGVTGKNGPVGNNPVTVKYKHIGSDDLAEB 384
QY 429 --NRIYDSEYEDVTNDDYAQLGNNDNNVINFNIDSP--YIIKVISKYDPNKKDYTTIQ 484
DB 385 VYAKLDDVSKFEDVTNDDSLDFDTGGYSLNENLQSKNYVIKGYDYSNASH-----439
QY 485 QVTWQTINNEYTGFRFA--SYDNITAFSTSSGGQGG-----DLPPEKTKIG 531
DB 440 --LEFQTHLFGYNYIYTSNLTKNGVAFYSNNAQCGDKDKLPIIEHSTPIELEPKSE 497
QY 532 DVVWEDVDKDGIONTNDNEKPLSNVLVTLT-----YPDGTSKSVPTDED 575
DB 498 PPVEKHELTGTIEESNDS-KPIDFEYHTAVEGAEGHAGST---IETED 542

RESULT 11
G90053
hypothetical protein fnb [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: G90053
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: G90053
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-961 <KUR>
A:Cross-references: GB:BA000018; PID:gl3702452; PIDN:BAB43593.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: fnb

Query Match 11.4%; Score 344.5; DB 2; Length 961;
Best Local Similarity 22.7%; Pred. No. 9.6e-08;
Matches 120; Conservative 125; Mismatches 203; Indels 81; Gaps 22;

QY 74 KESSVESNSIDTAQOPSHTTINREESVOTSDNVEDSHVDFANSKIKESNTESGKEE 133
DB 33 KEAASEQNNTIVBESGS-SATESKASETQTTNNV-----NTIDETQSYSA--- 78
QY 134 NTIQPNKVKEDSTTSQPSGY--TNIDKISNQDELMLNPINEYENK-----ARPLSTT 185
DB 79 TSTEQPSKSTQVTTBEAPTVAQPKVETEMKSDQLPSEKVADEKTTGTQVDIAQPSNV 138
QY 186 SAQPSIKR-----VTVNQLAAE---QGSNVNHLIKVTDQISITEGYDDSEGVIAKHAEN 236
DB 139 EIKPRMKASADVTASEKEVAEEAKATGTDVTKNKEVETESSL-ECHNKNQNVNPNHNAQR 197
QY 237 LIYDVTFEVDVKSGDPTMTVIDIKNTVPDLTDSFTIPKIDKNS-GEIATGTVDNKNKQIT 295
DB 198 VTLKYKMFGEKIGADYDFDTLSDNVETHGISLTKRKVPKESKTEKVMANGQVINE-R 256
QY 256 QITFTFTDYDKYENIKAHKLKLTSDYDKSKVPNNNTKLDVEYKTALESVNKTIIVFYORP 355
DB 257 TIRYFTFTDYINNKDLTAELNLFIDPTTQKNGO-KVEVILGQNKVSKFEIDIKYLDG 315
QY 356 NENR---TANLQSMFTNIDTKNH---TVEQTIYINPLRYSAKETNVAISNGSDGSTIID 409
DB 316 VKDRMGVTN-----GRIDLKKEGKFSHFAYVKPNQSL--TSVITGVQVTSYKQSA 368
QY 410 DSTIIKVKYKVDNQNLPS--NRIYDSEYEDVTNDDYAQLGNNDNNVINFNIDSP--PY 465
DB 369 NNPTVKYKHIGSDLAESVYAKLDDTSKPEDVTEKVNLSYTSNGGYTLNLGLDLSKDY 428
```

```
QY 466 IIKVISKYDPNKKDYTTIQOTVTMTQTTINEYTGEBRTASY-----DNTIATSTSS 515
DB 429 VIKYGEVDQNAKD-----LNFRTHLSGHYKYPYPPYPPYVQLTWNNGVAFYSNN 481
QY 516 GQGGG-DLPPEKTKIGDYVWEDVDKD-----GIQNTNDNEKPL 553
DB 482 AKGDKPKNPDIIEKSEPIDLDIKSEPPVPEKHELTGTIEESNDS-KPI 529

RESULT 12
H90053
hypothetical protein fnb [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: H90053
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: H90053
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1038 <KUR>
A:Cross-references: GB:BA000018; PID:gl3702453; PIDN:BAB43594.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: fnb

Query Match 10.8%; Score 326; DB 2; Length 1038;
Best Local Similarity 22.7%; Pred. No. 6.7e-07;
Matches 137; Conservative 109; Mismatches 235; Indels 122; Gaps 30;

QY 35 GIEKRSRSTESTNVDENAEATFLQKTPQDN-----THLTFEEYKESSESVSSSSIDTAQ 90
DB 30 GQDKEAASQKTTVEEN-----GNSATDNKTSQTATTATNVNHHIEETOSYNATV--TE 82
QY 91 QPSHTT-INREESVOTSDNVEDSHVDFANSKIKESNTESGKEENTIEQNKVKEDSTTS 149
DB 83 QPSNATQVTEEPKAVQ-----APQTAQPAENVETVKEE---EKP-QVKE---TT 125
QY 150 QPSGYTNIDKISNQDELMLNPINEYENK-----ARPLSTTSQPSKIKRVTVNQLAA 201
DB 126 QP-----QDNGSNQDVLTPKKVQNGQTQVQVAPQRTASESKPRVTR-SADVAEA 178
QY 202 EQGSNVNHLIKVTDQISITEGYDDSEGVIAK-----HDAENLIYDVTFEVDVKVSGD 253
DB 179 KEASDVSE-VKGTD--VTSKVTVESGSEIAPQGNKVEPHAGQRVLYKLPADGLKRGD 235
QY 254 TMTVDIDKNTVPDLTDSFTIPKIDKNS-----GEIATGTVDNKNKQITFTFDVVDKY 308
DB 236 YPDTLISNNVTVGVSTARKVPEIKNGSVVMATGEILNG-----NRYTFNIEHK 288
QY 309 ENIKAHKLKLTSDYDKSKVPNN-----NTKLDVEYKTALESVNKTIIVFYORPNNRANL 363
DB 289 VEVATANLEINLFDPTKTVQSGNQSGQKITSKLNGE-----ETEKTI PVVYNPVGSNSVTN 342
QY 364 QSMFTNIDTKNHITVEQTIYINPLRYSAKETNVAISNGSDGSTIIDSTIIKVKYK-VGDN 422
DB 343 NGSIETFNKESNKFTHIAVYKPMNGQSNST-VSVTGLTTEGNSLAGGQPTVKYVEYLKGG 401
QY 423 QNLPSDNRIY---DYSEYEDVTNDDYAQLG--NNNDVNINFGNIDSPYIIKVISKYDPN 476
DB 402 DELPQS--VYANTSDTNKFDVTKENNGKLSVQDNGSYSINLNDKLDKTVYIHYTGEVLQ 459
QY 477 KDDYTTIQOTVTMTQTTINEY-----GEFRTASVDNTIATSTSSGGQGGDLPPEK 526
DB 460 SDQ-----VNFRTELYGPERAYKSYVYVGGYR-LTWDNGLVLYSNKADGNG----- 505
QY 527 TYKIGDYVWE---DVDKDGIONTNDNEKPLSNVLVTLTTPDGTSKSV---RTDEGKYQ 579
```

[illegible]

Search completed: October 5, 2004, 20:53:07  
Job time : 25.2079 secs





```

FT REPEAT 745 782 FIBRONECTIN-BINDING DOMAIN.
FT REPEAT 783 820 D-1.
FT REPEAT 821 859 D-2.
FT REPEAT 860 878 D-3.
FT DOMAIN 879 948 D-4 (INCOMPLETE).
FT REPEAT 879 948 5 X TANDEM REPEATS, PRO-RICH (WR).
FT REPEAT 879 948 WR1.
FT REPEAT 893 906 WR2.
FT REPEAT 907 920 WR3.
FT REPEAT 921 934 WR4.
FT REPEAT 935 948 WR5.
FT SITE 985 986 LPXTG SORTING SIGNAL (POTENTIAL).
FT MOD_RES 985 985 AMIDG-LINKED TO CELL WALL (POTENTIAL).
SQ SEQUENCE 1018 AA; 111780 MW; 58175E0020E81F1F CRC64;

Query Match 10.0%; Score 300.5; DB 1; Length 1018;
Best Local Similarity 21.2%; Pred. No. 1.1e-05;
Matches 123; Conservative 118; Mismatches 243; Indels 97; Gaps 25;

QY 35 GIEKRSDETSNTVNDENATFLQKTPQDN---THLTBEVKSSVSSSSIDTAQ 90
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
30 QDKEAAASEQKTTVEEN---GNSATDNKTSQTATTNVNHNIEETQSYNATV--TE 82
QY 91 QPSHTT-INREE---SVQTSNVEDSHVDFPANSKIKESNTESKEENTIEQPNKVEDS 146
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
83 QPSNATQTTBEAKVAQAPQAPQAPANETVKEEVKVE-----EAKPQVKETT 130
QY 147 TTSQPSG---YTNIDEKISNODELNLPINEYENKARPLSTSAQPSIKRVT-----V 196
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
131 QSQDSNGDQDQVLTPKKATQNAQVETQV---EVAQPRTASESKPRVTSADVAEAKA 186
QY 197 NQLAAEGSNVNHILKVTQDSITEGYDDSEGVKAHDAENLIYDVTFEVDKVKSGDTMT 256
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
187 SNAKVETGTDVTSKVTVEIGSI-EGHNNTNKV-EPHAGQRAVLKYLKFKENGLHQGDYD 244
QY 257 VDIDKNTVPSDLTDSFTPIKONS-----GEIATGYDNKKNQIYTTDYDKYENI 311
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
245 FTLSNNVTGVSHTARKVPEIKGVSVMATGEVLEGG-----KIRYFTNDIEDKVDV 297
QY 312 KAHKLKTSYIDKSKVPNNNTKLDVEYKTAALSSV-----KTIITVEYORPENENTANLQSM 366
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
298 TAELEINLFIDPKVTQNGN-----QTITSTLNEEQTSKELDVKYKGIGNYANLGS 351
QY 367 FTNIDTQKHTVEQTIYINPLRYSAKETNVNINSGDEGTSIIDSTIIKVK-VGNQNL 425
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
352 IETFNKANNRFSHVAFIKP--NNGKTSVTVTGTLMKGSQNGQPKVRIFEYLGNNEDI 409
QY 426 PDSNRIY---DYSEYEDVTNDYQALG---NNNDVNIINFGNIDSPYIIKVISYDPNKO- 478
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
410 AKS--VVANTTDSKFKVETSNMGNLNLQNGSYSNIENLQKTYVHVHYGELYNGTDE 467
QY 479 -DYTTIQQTVMQTTINEYTGFEPTASNDNTIAFSTSSGOGGDLPP-----EKTYKIGD 532
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
468 VDERT-QMVGHPEQLKYKYDGRGYTLTWNDGLVLYSKANGNEKNGPIIQNNKFEYK--- 523
QY 533 YVNEVDVKDGIQNTNDNEKPLSNVLTLTYPDGTSKSVRTD 573
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
524 ---EDTIKETLTGQYDKN-----LVTTVBEEDYDSLTLDID 555

RESULT 2
ARP_PLAFA
ID ARP_PLAFA STANDARD; PRT; 537 AA.
AC P04931;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Asparagine-rich protein (AG319) (ARP) (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5633;
RN [1]
RP SEQUENCE FROM N.A.
```

```

RX MEDLINE=86176787; PubMed=2421257;
RA Stahl H.-D., Bianco A.E., Crewther P.E., Burkot T., Coppel R.L.,
RA Brown G.V., Anders R.F., Kemp D.J.;
RT "An asparagine-rich protein from blood stages of Plasmodium
RL falciparum shares determinants with sporozoites.";
RL Nucleic Acids Res. 14:3089-3102(1986).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; M24328; AAA29491.1; -.
PIR; A23770; A23770.
InterPro; IPR008928; Glyco_trans_6hp.
KW Malaria.
FT NON_TER 537 537
FT TER 537 537
SQ SEQUENCE 537 AA; 63000 MW; A6D76B15318CC239 CRC64;

Query Match 6.7%; Score 203; DB 1; Length 537;
Best Local Similarity 17.8%; Pred. No. 0.054;
Matches 93; Conservative 106; Mismatches 212; Indels 112; Gaps 19;

QY 11 NNNQSLTNDNNQIKKEETNNYDGEKESDRTSTTVNDENATFLQKTPQDNTHL 70
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
4 NKNNNNDGDN--INYQNTNEFKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNN 61
QY 71 EYKSSSVSSSSSID--TAQPSHTTT-----NRESVOT-----SDNVEDSHVSD 119
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
62 EHLR-NNSIDMNSNINNYTNQTRFSSPMENENENENKNNKNNKNNKNNKNNKNN 120
QY 120 SKIKSNTSGKEENTIEQPNKVKEDSTTSQPSGYTNIDEKISNODELNL-PINEY 178
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
121 SSMK--NTDNKNTDTSYNNKGTINNDN-----NNMDYLRNININNEYKGS 163
QY 179 ASPLSTSAQPSIKKVTNNQALAAEQSNVNHILKVTQDSITEGYDDSEGVKAHDAEN 238
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
164 AKKKFTN-----YNNKNNLKFQNNNDNNNNNNNNNNNNNNNNNNNNNNNNNN 217
QY 239 YDVTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFTPIKONSGBIATGTVDNKNKQIT 298
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
218 RNNINIKRNLN-----NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 268
QY 299 YTFDYDVKYENIKAHKLKTSYIDKSKVPNNNTKLDVEYKTAALSSV----- 345
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
269 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 328
QY 346 -KTIITVEYORPENENTANLQSM-----FTNIDTKNHTV-----EQTIYIN- 384
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
329 VQSMNNYNN-NESTANPNQNNFQETNNDNNKNNNNNNNNNNNNNNNNNNNNNN 387
QY 385 -----FLRYSAKETNVNI-----SGNGD-----EGSTIIDSTIIKVKYK- 420
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
388 PSRAVGYNVNNYLNNNNNNNNSAVNNNSNGNNNNNNNNNNNNNNNNNNNNNNNN 447
QY 421 DNQNLPSDNRIYDYSEY-----EDVTNDDYQALGNNDVNNIN 457
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
448 MNESINNNNTLNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 490

RESULT 3
PI4K_DICDI
ID PI4K_DICDI STANDARD; PRT; 1093 AA.
AC P54677;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phosphatidylinositol 4-kinase (EC 2.7.1.67) (PI4-kinase) (PtdIns-4-
kinase) (PI4K-alpha).
```

```
GN PIKD OR PIK4.
OS Dictyostelium discoideum (Slime mold).
OS Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RX MEDLINE=96009592; PubMed=7565716;
RA Zhou K., Takegawa K., Emr S.D., Firtel R.A.;
RT "A phosphatidylinositol (PI) kinase gene family in Dictyostelium
RT discoideum: biological roles of putative mammalian p110 and yeast
RT Vps34p PI 3-kinase homologs during growth and development.";
RL Mol. Cell. Biol. 15:5645-5656(1995).
CC -/- FUNCTION: Acts on phosphatidylinositol (PtdIns) in the first
CC committed step in the production of the second messenger
CC inositol-1,4,5,-trisphosphate.
CC -/- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1D-myo-inositol = ADP +
CC 1-phosphatidyl-1D-myo-inositol 4-phosphate.
CC -/- SIMILARITY: Belongs to the PI3/P14-kinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U23479; AAA85725.1; -.
DR PIR; T18275; T18275.
DR DictyBase; DDB0001907; pikD.
DR InterPro; IPR008938; ARM.
DR Pfam; PF00454; P13_P14_kinase; 1.
DR SMART; SM00146; P13Kc; 1.
DR PROSITE; PS00915; P13_4_KINASE_1; 1.
DR PROSITE; PS00916; P13_4_KINASE_2; 1.
DR PROSITE; PS02290; P13_4_KINASE_3; 1.
KW Transferase; Kinase; Multigene family.
FT DOMAIN 146 156 ASP-RICH.
FT DOMAIN 184 200 POLY-ASN.
FT DOMAIN 203 206 POLY-ASN.
FT DOMAIN 227 233 POLY-ASN.
FT DOMAIN 277 315 POLY-ASN.
FT DOMAIN 441 445 POLY-ASP.
FT DOMAIN 455 466 POLY-THR.
FT DOMAIN 494 501 POLY-GLY.
FT DOMAIN 690 696 POLY-THR.
FT DOMAIN 751 754 POLY-GLN.
FT DOMAIN 761 772 POLY-THR.
FT DOMAIN 775 785 POLY-THR.
FT DOMAIN 833 1093 P13K/PI4K.
SQ SEQUENCE 1093 AA; 123017 MW; 8762BC78355AA635 CRC64;

Query Match 6.7%; Score 202; DB 1; Length 1093;
Best Local Similarity 20.0%; Pred. No. 0.14;
Matches 116; Conservative 100; Mismatches 223; Indels 142; Gaps 24;

QY 1 SSDEKNDVNNQSIINTD--DNQIITKBEETNNYDGIKRESDRTSTTVD--ENEATPL 58
DB 184 NSNNNNNNNNNSNDNNNNNNNNNNELIPNENSDNSINDENQYGNNSNNNNISGNNNIKI 243
QY 59 QKTPQDNTLTHEEVKSESSVESNSSSI--DTAQQP-----SHTTINREESVQT 105
DB 244 DINSON-----KSDSNIELTSLTCLCETKTSPIKODMNNNNNNNNNNNNNNNN 294
QY 106 SDNVESHVDFPANSKESN-----TESGKENTIEQNKVKEDSTTSQPSG 153
DB 295 NNNNNNNNN--NNNNNNNNNYGHINGSLSLDGIQGYISQPNDFIENTITQLKRN 351
QY 154 ---YTIIDEKISNQDELLNLPINEYENKAPLSTSAQPSIKKVTYNQAAEGGSNNVHL 210
DB 352 RIIYKKVEKKEKELATRLRECFECSVITCSRPLITRPTSLPSPLISYNSGKIGGNTHKI 411
```

```
211 IKVTQOSTEGYDDSEGVKAHAENLIYDVTFFVDDKVKSGD-----TMTVDIDKNTVP 265
412 LSPSSVDSTSLISEDDKIIEKEEDNVVED---DDDEVNSEDFIPTATTATTTTTTIP 468
266 SLDLTSFTIPKIKDNGS---EIIATGTYDNKNKQIYTFYTDYVDKYENIKAHKLKLSYID 322
469 NHLSKTTSGVGINSSSTPININSAGAGGGEINHI-----GYDDI-----SYLD 515
323 KSKVNNNTKL--DVEYKTA-----LSSVKNKTITVE---YORPNENRTANLOS 365
516 KCKTPAESKLSHDHDFEPELSKSHRCDYLDNLSPIQKLAHISKILLPIPIDLRQAKLKH 575
366 MFTNIDTKNHTVEQIYINPLRYSAKETNNVNSGNGDEGSTIIDDSTIIKYKVGDNQNL 425
576 ---EISLLNINPLGLYV-PLWQSS-----NHHCVVRI-----P 605
426 PDSNRIYDSE-----YEDVTNDVYAQGNNDVNINFGNIDSPYIIKVIISKYDPNKDD 479
606 PREVKILNSRRVFPFLVLLEVIESEHALSSN-----IFEVVSY-----LQ 647
480 YTTIOQTVMQTTINEYTG-----FRTASYDNTIAFTSS 515
648 YTTGNSALKKDDIKRKYSEKFKKSLNSSINSTISNDS 688

RESULT 4
HWM1 MYCGE STANDARD; PRT; 1139 AA.
AC Q49413; Q49365;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytoadherence high molecular weight protein 1 (Cytoadherence accessory
DE protein 1).
GN HWM1 OR MG312.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2037;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bitt K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403(1995).
RN [2]
RP SEQUENCE OF 721-847 FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=94075230; PubMed=8253680;
RA Peterson S.N., Hu P.-C., Bitt K.F., Hutchison C.A. III;
RT "A survey of the Mycoplasma genitalium genome by using random
RT sequencing.";
RL J. Bacteriol. 175:7918-7930(1993).
CC -/- FUNCTION: COMPONENT OF THE CYTOSKELETON-LIKE STRUCTURE WHICH
CC STABILIZES THE SHAPE OF THE WALL-LESS MYCOPLASMA. THIS
CC CYTOSKELETON-LIKE NETWORK OF ACCESSORY PROTEINS CONTAINING HMW
CC PROTEINS 1 TO 5 ALLOWS THE PROPER ANCHORING OF CYTADHESIN PROTEINS
CC IN THE MYCOPLASMA MEMBRANE AT THE ATTACHMENT ORGANELLE (BY
CC SIMILARITY).
CC -/- SUBCELLULAR LOCATION: Localizes specifically to the attachment
CC membrane (By similarity).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC
```

CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC	EMBL; U39712; AAC71534.1; -
DR	EMBL; U02261; AAD12527.1; -
DR	PIR; B64234; E64234.
DR	TIGR; MG312; -
KW	Cycadherence; Structural protein; Complete proteome.
SQ	SEQUENCE 1139 AA; 130531 MW; 0011D328803DD856 CRC64;

<b>Query Match</b>					
Best Local Similarity 6.5%; Score 196.5; DB 1; Length 1139;					
Matches 145; Conservative 116; Mismatches 274; Indels 248; Gaps 34;					
Qy	1	SSDEKND-----VINNOSINTDNNQIIKKEETNNYDGIEKRESDTESTTN--VD 51			
Dd	402	NKDATNDSLSNEPFIQNSNETASDDVHYESKEPIHDY----KFGSDISQSNSNLSLE 457			
Qy	52	ENEATFLQKTPQDNTHL-----TEEEVKESSSESSNSSIDTAAQQSHHT 96			
Dd	458	SEPVKFSETAPD-AHPESQSEPVQVQYDIYQNEELKPTLDQPSSDDYP--AKQPTDEN 514			
Qy	97	INREE-----SVQTSNDVEDSHVSDFANSKIENSTESGKEENTIEQNKVKESTTSQPS 152			
Dd	515	YGFDNDLPPEVKOPESVVDQPPSSDDHHFAKQPESTIDSYSFSDLPQP-----TLDQPS 567			
Qy	153	-----GYNTIDEKISNODEL----- 167			
Dd	568	LDDHVQYNPDHHEELKPVASEQNYYGVGFDOVQOANLDNNEEIQTAEKVTTFDESQQAQ 627			
Qy	168	-----LNLPIN-EVENKARPLSTTSQAOPSIKRVTNVQLAEOGSNVNHLIK-VTDQSITE 220			
Dd	628	VDSYQLPIDTDQDQDTFFSSSFETOPTVEQFD-----QVNSEVNDQFKPEITKEPVL 681			
Qy	221	GYPDSEGGIKAHAENLIYDVTVFEVDKV-----KSGDT 254			
Dd	682	SFNKQDVVETSNYTNMLQKFEDIQSNKITITTKSSPOIPTTLPIFSVSNRIEYKPVT 741			
Qy	255	MTVDIDKN-----TVPSDLTDSFTPK-----IKNSGEIATG---TYDNKNKQIT 298			
Dd	742	LALDNKESQOEQITINSITEDSKTLAKTILSVQLQINSLNQSIIVTSVRLLDKDDQLT 801			
Qy	299	YTFTDVVDKYENIKAHLKTSYIDKSVPNNNTKLQVEYKTALSSVNK-----TIIVEVQ 353			
Dd	802	INTVNSEDOQPKLEVFWAKEPVEEHSITQN--KQSVEDKSELDNFNKKSDLYKIITSEIK 859			
Qy	354	RPNENRTANLQSMFTNIDTKNHVTVEGTIYNPL-----RYX 389			
Dd	860	RGLNPTINFDAIFOMNDYQMSVKQSFHLDNPTVTVYKNQISERYLI IKKELQSELSRLI 919			
Qy	390	AKETNNISGNDEGSTIIDSTLIK-----VTKVGDN-ONLPDSNRIVDY----- 434			
Dd	920	DQNELNVQFNNAKNTITLQKEEMIRSLASDFAIAYKPSNSEYQLOKSGEIMRHVQRALT 979			
Qy	435	-----SEYEDVTNDDYAOLGNH--NDVNINFGNIDSPYIIKVIISKYD--- 474			
Dd	980	ENEKKIESIGSGLKQLKTVVNSCCEITMKNINKLNTLAFACKERDPLL--LSNFDSVTD 1037			
Qy	475	-----PNK--DDYTTTQQVTVMQTTINEVGTGBFRFASYNNTIAFTSSQGOGDLPPEKT 527			
Dd	1038	NGLVEFPQLMDDB--LIDFNTFDNISNEQLDDFIYENMDRNIDFEPGFN----- 1085			
Qy	528	YKIGDVVWVEDVD--KDGQONTNDNEKPLSNVLATLTIPDGTGSK-SVRTDED-----GKY 578			
Dd	1086	---NDFV--DIDAKVWDMSAFSVNDLDETIV-----PDRTSNFSSLLEDLFESSGD 1135			
Qy	579	QFD 591			
Dd	1136	SLD 1138			

RESULT 5  
RBP2 PLAVB

ID	RB22_PLAIVE	STANDARD;	PRT; 2867 AA.
AC	Q00759; Q9N2M3;		
DT	01-APR-1993 (Rel. 25, Created)		
DT	15-MAR-2004 (Rel. 43, Last sequence update)		
DT	15-MAR-2004 (Rel. 43, Last annotation update)		
DE	Reticulocyte binding protein 2 precursor (PVRBP-2).		
GN	RB2-2 OR RBP2.		
OS	Plasmodium vivax (strain Belem).		
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.		
OX	NCBI_TaxID=31273;		
RN	[1]		
RP	SEQUENCE FROM N.A., AND REVISIONS TO 2438-2439.		
RX	MEDLINE=20299192; PubMed=10838229;		
RA	Galinski M.R.; Xu M.; Barnwell J.W.;		
RT	"Plasmodium vivax reticulocyte binding protein-2 (PVRBP-2) shares structural features with PVRBP-1 and the Plasmodium yoelii 235 kDa rhoptry protein family.";		
RT	Cell 69:1213-1226(1992).		
RL	Cell 69:1213-1226(1992).		
CC	-!- FUNCTION: Involved in reticulocyte adhesion. Specifically binds to human reticulocyte cells.		
CC	-!- SUBCELLULAR LOCATION: Membrane-bound (Probable).		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC	-----		
CC	EMBL; AF184623; AAF76525.1; -.		
DR	HSPSP; P03069; IGCM		
KW	Malaria; Receptor; Signal; Transmembrane; Repeat.		
FT	SIGNAL	1	POTENTIAL.
FT	CHAIN	22	2867 RETICULOCYTE BINDING PROTEIN 2.
FT	DOMAIN	22	2805 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	2806	2826 POTENTIAL.
FT	DOMAIN	2827	2867 CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	44	133 ASN-RICH.
FT	DOMAIN	560	758 LYS-RICH.
FT	DOMAIN	1112	1285 LYS-RICH.
FT	DOMAIN	2758	2785 7 X 4 AA TANDEM REPEATS OF H-D-D-T.
FT	REPEAT	2758	2761 1.
FT	REPEAT	2762	2765 2.
FT	REPEAT	2766	2769 3.
FT	REPEAT	2770	2773 4.
FT	REPEAT	2774	2777 5.
FT	REPEAT	2778	2781 6.
FT	REPEAT	2782	2785 7.
SQ	SEQUENCE	2867 AA; 331433 MW; 6E7D8CA71AFBFFD3 CRC64:	

Query Match	6.4%	Score 192.5;	DB 1;	Length 2867;
Best Local Similarity	22.3%;	Pred. No. 1.1;		
Matches 123;	Conservative 94;	Mismatches 192;	Indels 143;	Gaps 30;
QY	28	EETNNYDGIKRSDETESTTNVDE	---NEATFLQKTQDNTHTLE	-----E 71
		:::	:::	
Db	435	EHLKMPENLUSKSKRMLYSTFEKLEGLDLINKITLMGSEQSDLSIIADSEKIITKSAE		494
QY	72	EVKSESSVSSNSSIDTAQPSHTTINREBSVQTSNVEDSHVSDPAN	-----SKIKES	125
		:::	:::	
Db	495	SLINSSSEELAKYALDSNEKINEIKKNYDQNI	-----LKVREFINKSLGILTSVKGT	546
QY	126	N--TESGKE--ENTTIPQPNKVEDSTTSQPSGVTNIDEKISNQDELLNLPINEYENKARP		181
		:::	:::	
Db	547	SQUSESDQOQIEPKIEIKKKKKDI	-----LERGKEPINT--MNEIKKKKKS	591







QY 391 K---ETNVNIGSGDEGTTIIDSTIIKVKVGNQNLPSNRIDYSEYEDVTNDYQA 447  
 Db 1235 KHGFHTMNLHGCG-----VSRQSHNSHHGNRQD-----R 1266  
 QY 448 LGNNNDV-----NINFGNIDSPIIKVISKVD-----PNKDDVTYTIQOQVTVMQTTI 493  
 Db 1267 GGNISGNVLNMGKNNNNFNPNPSRY-----NLYDKKLDLDLYENRNDSTT-KELIKKLAEI 1320  
 QY 494 NEYTGFEFTASVDNTI 509  
 Db 1321 NKCENEISVKYCDHMI 1336

RESULT 10  
 LMA2 HUMAN  
 ID LMA2 HUMAN STANDARD; PRT; 3110 AA.  
 AC P24043; Q14736; Q93022;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Laminin alpha-2 chain precursor (Laminin M chain) (Merossin heavy chain).  
 GN LMA2 OR LAM.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=94124633; PubMed=8294519;  
 RA Vuolteenaho R., Niesinen M., Sainio K., Byers M., Eddy R.,  
 RA Hirvonen H., Shows T.B., Sariola H., Engvall E., Tryggvason K.;  
 RT "Human laminin M chain (merossin): complete primary structure,  
 RT chromosomal assignment, and expression of the M and A chain in human  
 RT fetal tissues.";  
 RL J Cell Biol. 124:381-394 (1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97066955; PubMed=8910357;  
 RA Zhang X., Vuolteenaho R., Tryggvason K.;  
 RT "Structure of the human laminin alpha2-chain gene (LMA2), which is  
 RT affected in congenital muscular dystrophy.";  
 RL J. Biol. Chem. 271:27664-27669 (1996).  
 RN [3]  
 RP SEQUENCE OF 1981-3110 FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Placenta;  
 RX MEDLINE=90238994; PubMed=2185464;  
 RA Ehrig K., Leivo I., Argraves W.S., Ruoslahti E., Engvall E.;  
 RT "Merossin, a tissue-specific basement membrane protein, is a  
 RT laminin-like protein.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:3264-3268 (1990).  
 RN [4]  
 RP VARIANTS GLN-545; HIS-619; LEU-919; HIS-2586 AND LYS-2614.  
 RA Panicker S.G., Mendell J.T., Chen L., Peng B., Sahenk Z.,  
 RA Marzluf G.A., Anato A.A., Mendell J.R.;  
 RT "Novel single base polymorphisms and rare sequence variants in  
 RT the laminin 2-chain coding region detected by RNA/SSCP analysis.";  
 RL Hum. Mutat. 13:174-174 (1999).  
 RN [5]  
 RP ERRATUM.  
 RA Panicker S.G., Mendell J.T., Chen L., Peng B., Sahenk Z.,  
 RA Marzluf G.A., Anato A.A., Mendell J.R.;  
 RL Hum. Mutat. 13:340-340 (1999).  
 RN [6]  
 RP VARIANT MDC1A PRO-2564.  
 RX MEDLINE=21476011; PubMed=11591858;  
 RA He Y., Jones K.J., Vignier N., Morgan G., Chevallay M., Barois A.,  
 RA Estournet-Mathaud B., Hori H., Mizuta T., Tome F.M.S., North K.N.,  
 RA Guicheney P.;  
 RT "Congenital muscular dystrophy with primary partial laminin alpha-2  
 RT chain deficiency: molecular study.";

Neurology 57:1319-1322 (2001).  
 [7]  
 RL VARIANTS MDC1A TYR-527 AND ARG-862.  
 RN MEDLINE=22439699; PubMed=12552556;  
 RX Tezak Z., Prandini P., Boscaro M., Marin A., Devaney J., Marino M.,  
 RA Fanin M., Trevisan C.P., Park J., Tyson W., Finkel R., Garcia C.,  
 RA Angelini C., Hoffman E.P., Pegoraro E.;  
 RT "Clinical and molecular study in congenital muscular dystrophy with  
 RT partial laminin alpha-2 (LMA2) deficiency.";  
 RL Hum. Mutat. 21:103-111 (2003).  
 CC -!- FUNCTION: Binding to cells via a high affinity receptor, laminin.  
 CC is thought to mediate the attachment, migration and organization  
 CC of cells into tissues during embryonic development by interacting  
 CC with other extracellular matrix components.  
 CC -!- SUBUNIT: Laminin is a complex glycoprotein, consisting of three  
 CC different polypeptide chains (alpha, beta, gamma), which are bound  
 CC to each other by disulfide bonds into a cross-shaped molecule  
 CC comprising one long and three short arms with globules at each  
 CC end. The alpha-2 chain is a subunit of laminin-2 (merossin) and  
 CC laminin-4 (S-merossin).  
 CC -!- SUBCELLULAR LOCATION: Extracellular; found in the basement  
 CC membranes (major component).  
 CC -!- TISSUE SPECIFICITY: Placenta, striated muscle, peripheral nerve,  
 CC cardiac muscle, pancreas, lung, spleen, kidney, adrenal gland,  
 CC skin, testis, meninges, choroid plexus, and some other regions of  
 CC the brain; not in liver, thymus and bone.  
 CC -!- DOMAIN: The alpha-helical domains I and II are thought to interact  
 CC with other laminin chains to form a coiled coil structure.  
 CC -!- DOMAIN: Domains VI, IV and G are globular.  
 CC -!- DISEASE: Defects in LMA2 are the cause of merossin-deficient  
 CC congenital muscular dystrophy type 1A (MDC1A) [MIM:607855]. MDC1A  
 CC is characterized by difficulty walking, hypotonia, proximal  
 CC weakness, hyporeflexia, and white matter hypodensity on MRI.  
 CC -!- SIMILARITY: Contains 1 laminin N-terminal domain.  
 CC -!- SIMILARITY: Contains 17 laminin EGF-like domains.  
 CC -!- SIMILARITY: Contains 2 laminin IV domains.  
 CC -!- SIMILARITY: Contains 5 laminin G-like domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z26653; CAAB1394.1; -;  
 DR EMBL; U66796; AAB18388.1; -;  
 DR EMBL; U66733; AAB18388.1; JOINED.  
 DR EMBL; U66734; AAB18388.1; JOINED.  
 DR EMBL; U66735; AAB18388.1; JOINED.  
 DR EMBL; U66736; AAB18388.1; JOINED.  
 DR EMBL; U66737; AAB18388.1; JOINED.  
 DR EMBL; U66738; AAB18388.1; JOINED.  
 DR EMBL; U66739; AAB18388.1; JOINED.  
 DR EMBL; U66740; AAB18388.1; JOINED.  
 DR EMBL; U66741; AAB18388.1; JOINED.  
 DR EMBL; U66742; AAB18388.1; JOINED.  
 DR EMBL; U66743; AAB18388.1; JOINED.  
 DR EMBL; U66745; AAB18388.1; JOINED.  
 DR EMBL; U66746; AAB18388.1; JOINED.  
 DR EMBL; U66747; AAB18388.1; JOINED.  
 DR EMBL; U66748; AAB18388.1; JOINED.  
 DR EMBL; U66749; AAB18388.1; JOINED.  
 DR EMBL; U66750; AAB18388.1; JOINED.  
 DR EMBL; U66751; AAB18388.1; JOINED.  
 DR EMBL; U66752; AAB18388.1; JOINED.  
 DR EMBL; U66753; AAB18388.1; JOINED.  
 DR EMBL; U66754; AAB18388.1; JOINED.  
 DR EMBL; U66755; AAB18388.1; JOINED.  
 DR EMBL; U66756; AAB18388.1; JOINED.  
 DR EMBL; U66757; AAB18388.1; JOINED.  
 DR EMBL; U66758; AAB18388.1; JOINED.



DR ENBL; U66759; AAB18388.1; JOINED.  
DR ENBL; U66760; AAB18388.1; JOINED.  
DR ENBL; U66761; AAB18388.1; JOINED.  
DR ENBL; U66762; AAB18388.1; JOINED.  
DR ENBL; U66763; AAB18388.1; JOINED.  
DR ENBL; U66764; AAB18388.1; JOINED.  
DR ENBL; U66765; AAB18388.1; JOINED.  
DR ENBL; U66766; AAB18388.1; JOINED.  
DR ENBL; U66767; AAB18388.1; JOINED.  
DR ENBL; U66768; AAB18388.1; JOINED.  
DR ENBL; U66769; AAB18388.1; JOINED.  
DR ENBL; U66770; AAB18388.1; JOINED.  
DR ENBL; U66771; AAB18388.1; JOINED.  
DR ENBL; U66772; AAB18388.1; JOINED.  
DR ENBL; U66773; AAB18388.1; JOINED.  
DR ENBL; U66774; AAB18388.1; JOINED.  
DR ENBL; U66775; AAB18388.1; JOINED.  
DR ENBL; U66776; AAB18388.1; JOINED.  
DR ENBL; U66777; AAB18388.1; JOINED.  
DR ENBL; U66778; AAB18388.1; JOINED.  
DR ENBL; U66779; AAB18388.1; JOINED.  
DR ENBL; U66780; AAB18388.1; JOINED.  
DR ENBL; U66781; AAB18388.1; JOINED.  
DR ENBL; U66782; AAB18388.1; JOINED.  
DR ENBL; U66783; AAB18388.1; JOINED.  
DR ENBL; U66784; AAB18388.1; JOINED.  
DR ENBL; U66785; AAB18388.1; JOINED.  
DR ENBL; U66786; AAB18388.1; JOINED.  
DR ENBL; U66787; AAB18388.1; JOINED.  
DR ENBL; U66788; AAB18388.1; JOINED.  
DR ENBL; U66789; AAB18388.1; JOINED.  
DR ENBL; U66790; AAB18388.1; JOINED.  
DR ENBL; U66791; AAB18388.1; JOINED.  
DR ENBL; U66792; AAB18388.1; JOINED.  
DR ENBL; U66793; AAB18388.1; JOINED.  
DR ENBL; U66794; AAB18388.1; JOINED.  
DR ENBL; U66795; AAB18388.1; JOINED.  
DR ENBL; M59832; AAB63215.1; -  
DR PIR; PX0082; M5HUMH.  
DR HSSP; Q60675; 1QUO.  
DR Genew; HGNC:6482; LAMA2.  
DR MIM; 156225; -  
DR MIM; 607855; -  
DR GO; GO:0005604; C:basement membrane; TAS.  
DR GO; GO:0005198; F:structural molecule activity; TAS.  
DR GO; GO:0007517; P:muscle development; TAS.  
DR InterPro; IPR008985; ConA like lec\_gl.  
DR InterPro; IPR006209; EGF like.  
DR InterPro; IPR008979; EGF bind like.  
DR InterPro; IPR000034; Laminin\_B.  
DR InterPro; IPR002049; Laminin\_EGF.  
DR InterPro; IPR001791; Laminin\_G.  
DR InterPro; IPR008211; LamNT.  
DR Pfam; PF00052; laminin\_B; 2.  
DR Pfam; PF00053; laminin\_EGF; 14.  
DR Pfam; PF00054; laminin\_G; 5.  
DR Pfam; PF00055; laminin\_Nterm; 1.  
DR PRINTS; PR00011; EGF\_LAMININ.  
DR ProDom; PD003031; Laminin\_B; 1.  
DR SMART; SM00180; EGF\_Lam; 15.  
DR SMART; SM00281; Lam5; 2.  
DR SMART; SM00282; LamG; 5.  
DR SMART; SM00136; LamNT; 1.  
DR PROSITE; PS00022; EGF\_1; 11.  
DR PROSITE; PS01186; EGF\_2; 3.  
DR PROSITE; PS01248; LAMININ TYPE EGF; 14.  
DR PROSITE; PS50025; LAM G DOMAIN; 5.  
KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;  
KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal; Polymorphism;  
KW Disease mutation.  
FT SIGNAL 1 22 POTENTIAL.

Query Match 6.2%; Score 186.5; DB 1; Length 3110;  
Best Local Similarity 19.9%; Pred. No. 2.1;

Matches 142; Conservative 102; Mismatches 256; Indels 215; Gaps 32;  
QY 21 NNQIIKKEET-----NNYDGIKSESDRTSTTNVDENEATFLOKTPQDNTLHTE 71  
DB NEKAIKLNELTGLTRDBAFERNLEGLQK-----EIDQMIKELRRKNLETKQIEAD 1742  
QY 72 EYKESSESSNSSIDTAQOPSHHTIN---RESVQTSNDVEDS-HVSPFANSKIKESN- 126  
DB ELVAAEALALKVKKLFGESEGENEEMKDLREKLDYKKNVDDAWDLLREATDKIHEARN 1802  
QY 127 -----TESGKE--ENTIEOPNKV-----KEDSTTS 149  
DB LPAVNQKMTALEKKKEAVESSCKRQIENTLKEGNDILDEANLADAINSIIYVEDIQIK 1862  
QY 150 QPSGYTNIDEKISN-----QDELLNLPINEYKNKARPLSTSA-----QPSIKRVTVNQL 199  
DB LPPMSEELNDKIDLSQEIKDRKLAEKVQSQAESHAQLNDSSAVLDGILDEAKNISFNAT 1922  
QY 200 AAEQG-SNVNHLIKVTDQSITEGYDDSEGVKAHAENLIYDVTPEVDDKVKSGDTMTVD 258  
DB AAFKAYSNIKQYI-----DEAEKVAK--BAKDLAHEAT-----KLATGPR---G 1961  
QY 259 IDKNTVPSDLTDSFTI-----PKIKNSGEIATGT-----YONKNKQITVTFDYVD 306  
DB LKEDAKGCLQKSFRLNEAKKLANDVKNEDHLNGLKTRIEADARNGDLLRTLNDTLG 2021  
QY 307 KYENI--KAHLKLTSYDKSKVPNNNTK-----LDVEYKLTALSSVNKTTIT 349  
DB KLSAIPNDTAAKLQAVKQKARQANDTAKDVLQAITELHQNLGDKVKNYKLNADSVAKTNA 2081  
QY 350 VEYQRPNEPR-TANLOSMTNID-----TKNHTVEQTIYINPLRYS 389  
DB V-KDPSKNKIADADATVKNLEQBADRLIDKLPKELEDNKLKNISEKELINOARKQ 2140  
QY 390 AKETNVNISGND-----EGS-----TIIDSTIIKV-----YK 418  
DB ANSIKVSVSSGGDCIPTYKPEIKKGSYNNIVNVKTAVADNLLFYLSAKFIDFLAEMR 2200  
QY 419 VGDNQNLPSNRIYDYSEYEDVTND-----YALQGNNDVNINFGNIDSPYIIKVIS 471  
DB KGVKSFMDVGVSGVGRVEYPLDTIDDSYWRIVASRTGRNGTISVR--ALDGPKASIVFS 2258  
QY 472 -KYDPNKDDYTTIQQVTMTQTTINEYGEFRFASVDNTIAFTSSGQGGDLPPEKTYKI 530  
DB THSTSPPGYTTILDVDANAMLFVGLTGKADAVRVITFTGCMGETYFDNKP-----I 2313  
QY 531 GDYVWEDVDKGIQNTNDNEKPLSNVLVTLTPD--GTSKSVRT-DEDGKYQFDG 582  
DB G--LWNFREKEG-----DCKGCTVSPQVEDSEGTIQFDG 2345

RESULT 11

TANA XENLA  
ID TANA XENLA STANDARD; PRT; 1744 AA.  
AC Q01550;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Tanabin.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Tadpole head;  
RX MEDLINE=92398961; PubMed=1524825;  
RA Hemmati-Brivanlou A., Mann R.W., Harland R.M.;  
RT "A protein expressed in the growth cones of embryonic vertebrate  
RT neurons defines a new class of intermediate filament protein.";  
RL Neuron 9:417-428(1992).  
CC -I- TISSUE SPECIFICITY: GROWTH CONES OF EMBRYONIC VERTEBRATE NEURONS.





```

Db 376 BEDSSSLCHGASLLKAIHTIYVNFVFSLNPSNQIAQATLQIISVVYDKIDLKQS----- 431
Qy 358 NRTANLOSMTFNIDTKNHTVETQTIYINPLYSAKETNNISGNGDEGSTIIDDSTIIKVV 417
Db 432 -----TSSVSLSTKHKHQQAELS-----EASENETAPL-----TLENMD 470
Qy 418 KVGONQNPDSNRIVDYSEYED--VTNDYYA 446
Db 471 KLND-----DEERLMDAQQPSIAITNQDLA 496

RESULT 14
RPB1_PLAFD
ID RPB1_PLAFD STANDARD; PRT; 2452 AA.
AC P14248;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE DNA-directed RNA polymerase II largest subunit (EC 2.7.7.6).
GN RPII.
OS Plasmodium falciparum (isolate CDC / Honduras).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5836;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90098832; PubMed=2690004;
RA Li W.B., Bzik D.J., Gu H., Tanaka M., Fox B.A., Inselburg J.;
RT "An enlarged largest subunit of Plasmodium falciparum RNA polymerase
RI II defines conserved and variable RNA polymerase domains.";
RL Nucleic Acids Res. 17:9621-9636(1989).
CC -!- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
CC of DNA into RNA using the four ribonucleoside triphosphates as
CC substrates.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -!- SUBUNIT: EACH CLASS OF RNA POLYMERASE IS ASSEMBLED FROM TEN TO
CC TWELVE DIFFERENT POLYPEPTIDES. THIS POLYPEPTIDE IS THE LARGEST
CC COMPONENT OF RNA POLYMERASE II.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- PTM: The tandem 7 residues repeats can be highly phosphorylated.
CC The phosphorylation activates POL2.
CC -!- MISCELLANEOUS: Three distinct zinc-containing RNA polymerases were
CC found in eukaryotic nuclei: polymerase I for the ribosomal RNA
CC precursor, polymerase II for the mRNA precursor, and polymerase
CC III for 5S and tRNA genes.
CC -!- SIMILARITY: Belongs to the RNA polymerase beta' chain family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X15561; CAA34560.1; -.
CC PIR; S07485; RNZQ2L.
CC DR InterPro; IPR007022; RNA_pol_A.
CC DR InterPro; IPR007080; RNA_pol_Rpb1_1.
CC DR InterPro; IPR007066; RNA_pol_Rpb1_3.
CC DR InterPro; IPR007083; RNA_pol_Rpb1_4.
CC DR InterPro; IPR007081; RNA_pol_Rpb1_5.
CC DR InterPro; IPR007075; RNA_pol_Rpb1_6.
CC DR InterPro; IPR007073; RNA_pol_Rpb1_7.
CC DR InterPro; IPR006592; RNA_pol_A_N.
CC DR InterPro; IPR00684; RNA_polII_repeat.
CC Pfam; PF04997; RNA_pol_Rpb1_1; 1.
CC Pfam; PF00623; RNA_pol_Rpb1_2; 1.
CC Pfam; PF04983; RNA_pol_Rpb1_3; 1.
CC Pfam; PF05000; RNA_pol_Rpb1_4; 1.
CC Pfam; PF04998; RNA_pol_Rpb1_5; 1.
CC Pfam; PF04992; RNA_pol_Rpb1_6; 1.

```

```

DR Pfam; PF04990; RNA_pol_Rpb1_7; 1.
DR Pfam; PF05001; RNA_pol_Rpb1_R; 12.
DR SMART; SM00663; RPOLA_N; 1.
DR PROSITE; PS00115; RNA_POL_II_REPEAT; 9.
KW Transferase; DNA-directed RNA polymerase; Transcription; Zinc; Repeat;
KW DNA-binding; Nuclear protein; Phosphorylation; Zinc-finger.
FT ZN FING 68 84 C2H2-TYPE (POTENTIAL).
FT DNA_BIND 378 411 POLY-ASN.
FT DOMAIN 707 725 LEUCINE-ZIPPER (POTENTIAL).
FT DOMAIN 1093 1128 POLY-ASN.
FT DOMAIN 1144 1159 VERY BASIC (POTENTIAL).
FT DNA_BIND 1182 1193 POLY-ASN.
FT DOMAIN 1258 1290 ASP/GLU-RICH (HIGHLY ACIDIC).
FT DOMAIN 1261 1290 6 X TANDEM REPEATS OF [YLV]-D(3,4).
FT DOMAIN 1602 1612 ASP/GLU-RICH (HIGHLY ACIDIC).
FT DOMAIN 1745 1759 ASP/GLU-RICH (HIGHLY ACIDIC).
FT DOMAIN 1806 1820 ASP/GLU-RICH (HIGHLY ACIDIC).
FT DOMAIN 2061 2246 HIGHLY DIVERGED HEPTAPEPTIDE REPEATS.
FT DOMAIN 2247 2384 CARBOXYL-TERMINAL 7-RESIDUE REPEATS.
SQ SEQUENCE 2452 AA; 278166 MW; F995E117F617A48F CRC64;

Query Match 6.1%; Score 183; DB 1; Length 2452;
Best Local Similarity 21.1%; Pred. No. 2.2;
Matches 145; Conservative 95; Mismatches 275; Indels 172; Gaps 28;

Qy 3 DEEKNDVINNOSINTDDNNQIIKKEETNNYDGIKRSSEDRTESTTNVDENEATFLQK-- 60
Db DDDLDDEYNSDININIGNRKYGNLTKNYD-----ENSMNPIDVVHKVNNFLEKLV 1336
Qy 61 -TPQDNTHITEVEKSSSVESNSISIDTAQPSHTTINRESVQTS-----DN 108
Db 1337 IIKQINSN-----DTLSVEAQNNATILLKAHLRTYLSKLLTOTHKVSVKGLDWLQ 1389
Qy 109 VE-----DSHVSDFANSKIKESNTES-----GKEENTIEQP-----NKV 142
Db 1390 IEKIFKYSCHPGECVGAALAAQISGEPAOTMLNTPHFGVGSKNVTGVPRLKELINIV 1449
Qy 143 KEDSTTSQPSGYTNIDEKISNQDELLNLPINENKAREPLSTTSAPSKRVTVNQLAAE 202
Db 1450 KNVKT---ESTIYLDMDVSNQOKAKDILTLEVTTLQSLTSHAQIIYDPNNTTILEE 1506
Qy 203 QSNVNHLLKVTQSDITEGYDSSGVIKAHDAENLIYDVFVDDKVKSGDWTVDIDKN 262
Db 1507 DKSWNVEFYEPFDDDTQ-YSLGEWVLR-----IQLTNIHVNEK-KLTMEKIVYIYS 1557
Qy 263 TVPSDLDTFTPIKIDNSGEII-----ATGYDNKNKQITFTDVTVDVYENIKAH 315
Db 1558 VFSSDELD---LIYTDNSEDIVLRIRVKYLNGEFNFMVYVDNANEQVDSEDEEHL 1614
Qy 316 -----KLTSYIDKSKVPNN-----NTK-----LDVEYKTALS 342
Db 1615 VANDRGNYDETKNSTPHHDYNNNTNIFKSKVKNMISSDINTKNEDSISINSSNNEQVK 1674
Qy 343 SVNKITVEYQPNENRANLOSMTNIDTKNHTVE-----QTIYINPLRYS 389
Db 1675 NINSPVSNMFMNNNNNNNNNDSSNINDIKVKNIKKEDGNEGALRGGGSDNTSALFGNKS 1734
Qy 390 AKETNV--NISGNGEGSTIIDDSITIIKVYKVGDNQNLPSNRIVDYSEYEDVTND-- 444
Db 1735 QKEDNIVNNNDNDDDD---DEEBEEDFLPGDHNVSPKNTKDGKNKNNKNNNNENK 1790
Qy 445 YAQLGNNNDVINFP---GNIDSPYIIKVISKVDPNKDDYTTTIOQTVMQTTTINEYTGEP 501
Db 1791 NKKSGNNNSNNTYDDGDVDN-----DNDDDDNDKNSDI-----TIKE----- 1829
Qy 502 TASYDNTAF-STSSGQGGDLPPKTYKIGDYVWEDVDKGIQNT---NDNEKPLSNVL 557
Db 1830 -----DNDVAFMKTSTKNAEEDL-----ELKNKNHSHISREDETFILKKMEQCLSTL 1881
Qy 558 -----VTILTPDGTSGSVRTDEDGKY 578
Db 1882 LRGIENITKVYMEESKITYSDNCKP 1908

```

RESULT 15  
 ID1\_MRC1\_YEAST STANDARD; PRT; 1096 AA.  
 AC P25589: P27513; P87003; Q07218; Q8NIN2;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE DNA replication checkpoint mediator MRC1.  
 GN MRC1 OR YCL061C OR YCL61C/YCL60C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 [1]  
 SEQUENCE FROM N.A.  
 RP STRAIN=S288C;  
 RX MEDLINE=92244356; PubMed=1574125;  
 RA Oliver S.G., van der Aart Q.J.M., Agostoni-Carbone M.L., Aigle M.,  
 RA Alberghina L., Alexandraki D., Antoine G., Anwar R., Ballesta J.P.G.,  
 RA Benit P., Berben G., Bergantino E., Biteau N., Bolle P.A.,  
 RA Bolotin-Fukuhara M., Brown A.J.P., Brown R., Buhler J.-M.,  
 RA Carignani G., Chanet R., Contreras R., Crouzet M., Daignan-Fornier B.,  
 RA De Haan M., Defoor E., Delgado M.D., Demolder J., Doira C., Dubois E.,  
 RA Dujon B., Duesterhoeft A., Erdmann D., Esteban M., Fabre F.,  
 RA Fairhead C.A., Faye G., Feldmann H., Fiers W.,  
 RA Francinques-Gaillard M.-C., Franco L., Frontali L., Fukuhara H.,  
 RA Fuller L.J., Gent M.E., Gigot D., Gilliquet V., Glandsdorff N.,  
 RA Goffeau A., Grenson M., Grisanti P., Grivell L.A., Haasemann M.,  
 RA Hatat D., Hegemann J.H., Herbert C.J., Hilger F., Hohmann S.,  
 RA Hollenberg C.P., Huse K., Iborra F., Indge K.J., Isono K., Jackman P.,  
 RA Jacq C., Jaquet M., James C.M., Jauniaux J.-C., Jia Y., Jimenez A.,  
 RA Kleinhaus U., Kreisel P., Lafranchi G., Lewis C., van der Linden C.G.,  
 RA Lucchini G., Lutzenkirchen K., Maat C.C., Mannheim G., Manzano M.E.,  
 RA Martignani E., Mathieu A., Maurer C.T.C., McConnell D., McKee R.A.,  
 RA Messenguy F., Mewes H.-W., Molemans F., Montague M.A., Navas L.,  
 RA Newton C.S., Olson M.V., Pallier C., Panzeri L., Pearson B.M.,  
 RA Perea J., Philippsen P., Piarard A., Planta R.J., Plevani P.,  
 RA Poetsch B., Pohl F.M., Purnelle B., Ramezani Rad M., Rasmussen S.W.,  
 RA Raynal A., Remacha M., Richterich P., Roberts A.B., Rodriguez F.,  
 RA Sanz E., Schaaff-Gerstenschlaeger I., Scherens B., Schweitzer B.,  
 RA Shu Y., Skala J., Slonimski P.P., Sor F., Soustelle C.,  
 RA Spiegelsberg R., Staveva L.I., Steensma H.Y., Steiner S., Thierry A.,  
 RA Thiores G., Triano L.N., Urrestazu L.A., Valle G., Vetter I.,  
 RA van Vliet-Reedijk J.C., Volckaert G., Vreken P., Warwington J.R.,  
 RA von Wettstein D., Wicksteed B.L., Wilson C., Wurst H., Xu G.,  
 RA Zimmermann F.K., Sgouras J.G.  
 RT "The complete DNA sequence of yeast chromosome III."  
 RL Nature 357:38-46(1992).  
 [2]  
 RP REVISIONS.  
 RA Gromadka R.;  
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP REVISIONS.  
 RA Valles G., Volckaerts G.;  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE OF 660-839 FROM N.A.  
 RC STRAIN=ATCC 28383 / FL100;  
 RX MEDLINE=90384830; PubMed=2169608;  
 RA Kern L.;  
 RT "The URK1 gene of Saccharomyces cerevisiae encoding uridine kinase."  
 RL Nucleic Acids Res. 18:5279-5279(1990).  
 [5]  
 RP FUNCTION, AND PHOSPHORYLATION.  
 RX MEDLINE=21571912; PubMed=11715016;  
 RA Alcasabas A.A., Osborn A.J., Bachant J., Hu F., Werler P.J.,  
 RA Bousset K., Furuya K., Diffley J.F., Carr A.M., Elledge S.J.;  
 RT "Mrc1 transduces signals of DNA replication stress to activate  
 Rad53."  
 RL Nat. Cell Biol. 3:958-965(2001).

[6]  
 FUNCTION, AND SUBCELLULAR LOCATION.  
 RX MEDLINE=22747592; PubMed=12865299;  
 RA Osborn A.J., Elledge S.J.;  
 RT "Mrc1 is a replication fork component whose phosphorylation in  
 response to DNA replication stress activates Rad53."  
 RL Genes Dev. 17:1755-1767(2003).  
 [7]  
 INTERACTION WITH CDC45.  
 RX MEDLINE=22825747; PubMed=12944972;  
 RA Katou Y., Kanoh Y., Bando M., Noguchi H., Tanaka H., Ashikari T.,  
 RA Sugimoto K., Shirahige K.;  
 RT "S-phase checkpoint proteins Tof1 and Mrc1 form a stable  
 replication-pausing complex."  
 RL Nature 424:1078-1083(2003).  
 CC -!- FUNCTION: Required for normal DNA replication. Phosphorylated in  
 response to DNA replication stress. Phosphorylation allows it to  
 mediate the activation of RAD53.  
 CC -!- SUBUNIT: Interacts with CDC45 in S phase.  
 CC -!- SUBCELLULAR LOCATION: Nuclear; associated with chromatin during S  
 phase.  
 CC -!- PTM: Phosphorylated by MEC1 and RAD53.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X59720; CAC42953.1; -;  
 DR EMBL; X53998; CAA37945.1; -;  
 DR PIR; S74279; S74279.  
 DR Germline; 138895; -;  
 DR SGD; S0000566; MRC1.  
 DR CO; GO:0006347; P:chromatin silencing at HML and HMR (sensu S. .; IGI.  
 DR GO; GO:0006348; P:chromatin silencing at telomere; IGI.  
 DR GO; GO:0000076; P:DNA replication checkpoint; IGI.  
 KW Nuclear protein; DNA replication; Phosphorylation; Coiled coil.  
 FT DOMAIN 488 542 COILED COIL (POTENTIAL).  
 FT DOMAIN 652 716 COILED COIL (POTENTIAL).  
 FT DOMAIN 748 808 L -> V (IN REF. 4).  
 FT CONFLICT 748 808 MISSING (IN REF. 4).  
 FT CONFLICT 808 808 MISSING (IN REF. 4).  
 SQ SEQUENCE 1096 AA; 124325 MW; 378345E503FFA81 CRC64;  
 Query Match 6.1%; Score 182.5; DB 1; Length 1096;  
 Best Local Similarity 18.9%; Pred. No. 0.88;  
 Matches 121; Conservative 117; Mismatches 234; Indels 169; Gaps 29;  
 QY 5 EKNDVINNQSIINTDNNQ-----IIKKEETNNYDGIKESDRTSTNNVDNEATFLQK 60  
 DB 534 EENDFQLNAHDSGSDSGSSGFGALSGNEIADYESSGSENDNRNRSDEKEDDEILKQK 593  
 QY 61 TPQDNTHL-----TEEEVKSSSVSESSNSIDTAQPSHTTN-----REESVQSDNV 109  
 DB 594 KSHHVXHHINESDSDETEVAKPKKAKDES-----LPKRIALNIGHYDNGTGEDTKF 645  
 QY 110 EDSHVSDPANSKIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQ---DE 166  
 DB 646 QETNVLDTONIE-----EVMALERTIE--NEVKDDVYNEEA-----DEAIRQLIDKE 692  
 QY 167 LNLPLINEYENKARPLSTTSQAQSIKRVTVN--QLAAEQSGSNVNHILIKVTQDSITGEYDD 224  
 DB 693 KLQKQKEKEHEAK-----IKELKKRGVTNPFMEAESESEDEWHGIGGAGGSDSDYDS 746  
 QY 225 -----SEGVIKAHDAENLIYDTFFVDKVKSGDWTVDID-----KNTVPS 266  
 DB 747 DLEKMIDDYSKNFNPHREMLAENKMD--IKMINKILYDIKNGGFRNKAKNSLEL 804  
 QY 267 DLTDSTFTPIPKDNGSGEIIATGYDNKNKQITVTFDYVDYKVENIKAHKLKLTSVIDKSKV 326  
 DB 805 ELSD-----DDEDDVLQ--QYRLKRLRLMR-----KRLLEIG---DDAKL 839

Search completed: October 5, 2004, 20:48:56  
Job time : 27.1685 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:41:22 ; Search time 65.362 Seconds  
(without alignments)  
2809.455 Million cell updates/sec

Title: US-10-806-288-13  
Perfect score: 3013  
Sequence: 1 SDEEKNDVNNQSIINTDD.....PDGTSKSVRTDEGKYQFDG 582

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues  
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3013	100.0	1056	16	Q8CQ72
2	3013	100.0	1092	2	O70022
3	2810.5	93.3	931	2	Q9K113
4	1160.5	38.5	1171	2	Q9KWX6
5	1144.5	38.0	1166	2	O86489
6	1143.5	38.0	1141	16	Q8NXX5
7	1137.5	37.8	1141	16	Q99W46
8	1137.5	37.8	1141	16	Q932F7
9	576	19.1	1315	2	O86488
10	567	18.8	1347	16	Q8NXX6
11	559	18.6	1385	16	Q99W47
12	499	16.6	1733	2	Q9K114
13	497	16.5	1633	16	Q8CWP4
14	489	16.2	953	16	Q99W48
15	485.5	16.1	1893	16	Q8KWH1
16	470.5	15.6	946	16	Q8NXX1

17	465	15.4	955	16	Q8NXX7	Q8nxx7 staphylococ
18	460.5	15.3	881	2	Q93MH7	Q93mh7 staphylococ
19	458.5	15.2	566	2	Q8KR22	Q8kr22 staphylococ
20	455	15.1	947	2	O86487	O86487 staphylococ
21	452	15.0	933	2	Q53653	Q53653 staphylococ
22	452	15.0	935	16	Q932C5	Q932c5 staphylococ
23	452	15.0	989	16	Q99VJ4	Q99vj4 staphylococ
24	397.5	13.2	877	16	Q99R07	Q99r07 staphylococ
25	396	13.1	913	2	O86476	O86476 staphylococ
26	384.5	12.8	907	16	Q8NUL0	Q8nul0 staphylococ
27	376	12.5	940	2	Q53682	Q53682 staphylococ
28	374.5	12.4	943	16	Q8NUU8	Q8nuu8 staphylococ
29	344.5	11.4	961	16	Q99RD3	Q99rd3 staphylococ
30	326.5	10.8	1015	16	Q8NUU7	Q8nuu7 staphylococ
31	326	10.8	1038	16	Q99RD2	Q99rd2 staphylococ
32	293.5	9.7	970	16	Q8DYL7	Q8dy17 streptococ
33	293.5	9.7	1310	16	Q8E473	Q8e473 streptococ
34	276	9.2	2616	5	Q8IIG1	Q8iig1 plasmodium
35	268.5	8.9	1301	5	Q8WSK5	Q8wak5 plasmodium
36	266.5	8.8	1160	2	Q8RM86	Q8rm86 streptococ
37	266.5	8.8	1301	5	Q8IHQ2	Q8ihq2 plasmodium
38	266	8.8	2402	2	Q9AER7	Q9aer7 staphylococ
39	265	8.8	3452	5	Q8IEA3	Q8iea3 plasmodium
40	261	8.7	1161	2	Q9X3M7	Q9x3m7 streptococ
41	261	8.7	2153	5	Q8ICM9	Q8icm9 plasmodium
42	259	8.6	1161	2	Q8RJI0	Q8rjl0 streptococ
43	253.5	8.4	3078	5	Q8IKH4	Q8ikh4 plasmodium
44	253.5	8.4	3207	5	Q8IHQ0	Q8ihq0 plasmodium
45	249.5	8.3	1859	5	Q8IC27	Q8ic27 plasmodium

ALIGNMENTS

RESULT 1  
Q8CQ72 PRELIMINARY; PRT; 1056 AA.  
AC Q8CQ72;  
DT 01-MAR-2003 (Tremblrel. 23, Created)  
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding protein.  
GN SE0331.  
OS Staphylococcus epidermidis.  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=1282;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 12228;  
RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z., Chen Z., Wen Y.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; A5016745; AAC03928.1; -;  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR008454; Cna\_B.  
DR InterPro; IPR001899; Gram\_pos\_anchor.  
DR Pfam; PF05738; Cna\_B; 2.  
DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
DR TIGRPFAM; TIGR01167; LPXTG\_anchor; 1.  
DR TIGRPFAM; TIGR01168; YSIRK\_signal; 1.  
DR PROSITE; P550847; GRAM\_POS\_ANCHORING; 1.  
KW Complete proteome.

SQ SEQUENCE 1056 AA; 115728 MW; 015869A9B5CA2723 CRC64;

Query Match 100.0%; Score 3013; DB 16; Length 1056;  
Best Local Similarity 100.0%; Pred. No. 2.5e-108;  
Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SDEEKNDVNNQSIINTDDNQIIKKETNNYDGIKESEDETESTTNNVDENATFLOK 60  
|||||

```
Db 75 SSDEKNDVNNQSIINTDDNNQIIKKEETNNYDGIKESDRTESTTNVDENATFLQK 134
Qy 61 TPQDNTHLTHEEVKESSESSVSSNSIDTAQPSHTTINREESVQTSNDVEDSHVDFANS 120
Db 135 TPQDNTHLTHEEVKESSESSVSSNSIDTAQPSHTTINREESVQTSNDVEDSHVDFANS 194
Qy 121 KIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQDELLNLPINEYENKAR 180
Db 195 KIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQDELLNLPINEYENKAR 254
Qy 181 PLSTTSQAQPSIKRVTVNQLAAEQGSNNVHLIKVTDQSIITEGYDDSEGVKAHDAENLIYD 240
Db 255 PLSTTSQAQPSIKRVTVNQLAAEQGSNNVHLIKVTDQSIITEGYDDSEGVKAHDAENLIYD 314
Qy 241 VTFEVDKVKSGDGMTVDIDKNTVPSDLTDSFTIPKIKDNGSGEIIATGTYDNNKQIITYT 300
Db 315 VTFEVDKVKSGDGMTVDIDKNTVPSDLTDSFTIPKIKDNGSGEIIATGTYDNNKQIITYT 374
Qy 301 FTDYVDKYENIKAHKLKLTYSIDKSKVPNNNTKLDVEYKKTALSSVKNKTIITVEYQRPNEKRT 360
Db 375 FTDYVDKYENIKAHKLKLTYSIDKSKVPNNNTKLDVEYKKTALSSVKNKTIITVEYQRPNEKRT 434
Qy 361 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVNSISGDEGSTIIDSTIIKVKYKVG 420
Db 435 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVNSISGDEGSTIIDSTIIKVKYKVG 494
Qy 421 DNQNLPSNRIYDYSEYEDVTNDYDIAQLGNNNDVNFNGNIDSPYIIKVISKYDPNKDDY 480
Db 495 DNQNLPSNRIYDYSEYEDVTNDYDIAQLGNNNDVNFNGNIDSPYIIKVISKYDPNKDDY 554
Qy 481 TTIQQTVMQTTINEYTGFEFTASVDNTIAFSTSSGQGGDLPPKTYKIGDYVWEDVDK 540
Db 555 TTIQQTVMQTTINEYTGFEFTASVDNTIAFSTSSGQGGDLPPKTYKIGDYVWEDVDK 614
Qy 541 DGIQNTNDNEKPLSNVLTLTYPDGTSKSVRTDEDEGKYQFDG 582
Db 615 DGIQNTNDNEKPLSNVLTLTYPDGTSKSVRTDEDEGKYQFDG 656

RESULT 2
O70022 PRELIMINARY; PRT; 1092 AA.
AC STRAIN=HB;
RX MEDLINE=98261511; PubMed=956732;
RA Nilsson M., Frykberg L., Flock J.I., Pei L., Lindberg M., Guss B.;
RT "A Fibrinogen-binding protein of Staphylococcus epidermidis.";
RL Infect. Immun. 66:2666-2673(1998).
CC -|- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY
AN AMIDE BOND (BY SIMILARITY).
CC EMBL; Y17116; CAA76638.1; -.
DR PIR; T30214; T30214.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR008454; Cna_B.
DR InterPro; IPR005877; Gpos_YSRK.
DR Pfam; PF05738; Cna_B; 2. Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; YsirK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; YsirK_signal; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor; Signal.
```

```
FT SIGNAL 1 51 POTENTIAL.
FT CHAIN 52 1092 FIBRINOGEN-BINDING PROTEIN.
SQ SEQUENCE 1092 AA; 119292 MW; 6542BC39AAD8B984 CRC64;

Query Match 100.0%; Score 3013; DB 2; Length 1092;
Best local similarity 100.0%; Pred. No. 2.6e-108;
Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSDEKNDVNNQSIINTDDNNQIIKKEETNNYDGIKESDRTESTTNVDENATFLQK 60
Db 75 SSDEKNDVNNQSIINTDDNNQIIKKEETNNYDGIKESDRTESTTNVDENATFLQK 134
Qy 61 TPQDNTHLTHEEVKESSESSVSSNSIDTAQPSHTTINREESVQTSNDVEDSHVDFANS 120
Db 135 TPQDNTHLTHEEVKESSESSVSSNSIDTAQPSHTTINREESVQTSNDVEDSHVDFANS 194
Qy 121 KIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQDELLNLPINEYENKAR 180
Db 195 KIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQDELLNLPINEYENKAR 254
Qy 181 PLSTTSQAQPSIKRVTVNQLAAEQGSNNVHLIKVTDQSIITEGYDDSEGVKAHDAENLIYD 240
Db 255 PLSTTSQAQPSIKRVTVNQLAAEQGSNNVHLIKVTDQSIITEGYDDSEGVKAHDAENLIYD 314
Qy 241 VTFEVDKVKSGDGMTVDIDKNTVPSDLTDSFTIPKIKDNGSGEIIATGTYDNNKQIITYT 300
Db 315 VTFEVDKVKSGDGMTVDIDKNTVPSDLTDSFTIPKIKDNGSGEIIATGTYDNNKQIITYT 374
Qy 301 FTDYVDKYENIKAHKLKLTYSIDKSKVPNNNTKLDVEYKKTALSSVKNKTIITVEYQRPNEKRT 360
Db 375 FTDYVDKYENIKAHKLKLTYSIDKSKVPNNNTKLDVEYKKTALSSVKNKTIITVEYQRPNEKRT 434
Qy 361 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVNSISGDEGSTIIDSTIIKVKYKVG 420
Db 435 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVNSISGDEGSTIIDSTIIKVKYKVG 494
Qy 421 DNQNLPSNRIYDYSEYEDVTNDYDIAQLGNNNDVNFNGNIDSPYIIKVISKYDPNKDDY 480
Db 495 DNQNLPSNRIYDYSEYEDVTNDYDIAQLGNNNDVNFNGNIDSPYIIKVISKYDPNKDDY 554
Qy 481 TTIQQTVMQTTINEYTGFEFTASVDNTIAFSTSSGQGGDLPPKTYKIGDYVWEDVDK 540
Db 555 TTIQQTVMQTTINEYTGFEFTASVDNTIAFSTSSGQGGDLPPKTYKIGDYVWEDVDK 614
Qy 541 DGIQNTNDNEKPLSNVLTLTYPDGTSKSVRTDEDEGKYQFDG 582
Db 615 DGIQNTNDNEKPLSNVLTLTYPDGTSKSVRTDEDEGKYQFDG 656

RESULT 3
Q9K113 PRELIMINARY; PRT; 931 AA.
ID O9K113
AC O9K113;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fibrinogen-binding protein sdrg.
GN SDRG.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K28;
RX MEDLINE=20340957; PubMed=10878118;
RA McCrea K.W., Hartford O., Davis S., Ni Eidhin D., Lina G.,
RA Speziale P., Foster T.J., Hook M.;
RT "The serine-aspartate repeat (Sdr) protein family in Staphylococcus
epidermidis.";
RL Microbiology 146:1535-1546(2000).
CC -|- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY
AN AMIDE BOND (BY SIMILARITY).
CC EMBL; AF245042; AAF72510.1; -.
DR
```



```
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR008454; Cna_B.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF05738; Cna_B; 2; Gaps 1.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
DR PROSITE; PS0847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 931 AA; 10295 MW; 591E657D97027116 CRC64;

Query Match 93.3%; Score 2810.5; DB 2; Length 931;
Best Local Similarity 93.6%; Pred. No. 1.3e-100;
Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;

QY 1 SSDEEKNDVNNQSIINTDDNQIIKKEETNNYDGIKRSDETESTTNVDENEATFLOK 60
DB 74 SSNEEKNDVNNQSIINTDDNQ--IKKEETNSNDAIENRSKDIQTSTTNVDENEATFLOK 132
QY 61 TPQDNTLHTEEEVKESSVSSNSIDTAQPSHTTINREESVQTSNDVEDSHVDFPANS 120
DB 133 TPQDNTQLKEEVKESPSSVSSNSMDTAQPSHTTINSEASIQTSNDNEENSRVDFPANS 192
QY 121 KIKESNTESGKEENTIEQPNKVEDSTTSQPSGYTIDEXISNQDELLNLPINEYENKAR 180
DB 193 KIIESNTESNKEENTIEQPNKVEDSITSQPSYKNIIDEXISNQDELLNLPINEYENKVR 252
QY 181 PLSTTSAQPSIKRVTVNQLAABQGSNNHLIKVTDQSITEGYDDSEGVKAHDAENLIYD 240
DB 253 PLSTTSAQPSKRVTVNQLAABQGSNNHLIKVTDQSITEGYDDSGIKAHDAENLIYD 312
QY 241 VTFEVDKVKSGDGTMTVDIDKNTVPDLTDSFTPIKIKNSGELIATGYDKNKQIITYT 300
DB 313 VTFEVDKVKSGDGTMTVDIDKNTVPDLTDSFPAIPKIKNSGELIATGYDNTNKQIITYT 372
QY 301 FTDYVDKYENIKAHKLTSYIDSKVPNNNTKLDVEYKTALSSVKNKTIITVEYQRPNEKRT 360
DB 373 FTDYVDKYENIKAHKLTSYIDSKVPNNNTKLDVEYKTALSSVKNKTIITVEYQKPNERT 432
QY 361 ANLQSMFTNIDTKNHTVEQTIYINPLAYSARETNVAINSGNDEGSTIIDDSTIIKYKVG 420
DB 433 ANLQSMFTNIDTKNHTVEQTIYINPLAYSARETNVAINSGNDEGSTIIDDSTIIKYKVG 492
QY 421 DNQNLPSNRIYDYSEYEDVTNDYAOLGNNNDVNFNGNIDSPYIIKIVLSKYDPNKDDY 480
DB 493 DNQNLPSNRIYDYSEYEDVTNDYAOLGNNNDVNFNGNIDSPYIIKIVLSKYDPNKDDY 552
QY 481 TTIQQTVMQTTINEYTGERTASYDNTIAFSTSSGGQGGDLPPKTYKIGDYVWEDVDK 540
DB 553 TTIQQTVMQTTINEYTGERTASYDNTIAFSTSSGGQGGDLPPKTYKIGDYVWEDVDK 612
QY 541 DGIONTNDNEKPLSNVLVTIYDPDGTSKSVRTDEGDKYQFDG 582
DB 613 DGIONTNDNEKPLSNVLVTIYDPDGTSKSVRTDEGDKYQFDG 654

RESULT 4
Q9KWX6 PRELIMINARY; PRT; 1171 AA.
AC Q9KWX6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Bone sialoprotein-binding protein.
GN BBP.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=024;
RX MEDLINE=20115096; PubMed=10642520;
RA Tung H.S., Guss B., Hellman U., Persson L., Rubin K., Ryden C.;
RT "A bone sialoprotein-binding protein from Staphylococcus aureus: a
RL Biochem. J. 345:611-619(2000).
CC 1- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY
EMBL; Y18653; CAB75732.1; -.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR008454; Cna_B.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF05738; Cna_B; 3.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
DR PROSITE; PS0847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 1171 AA; 127123 MW; CSBC812F9DA5A884 CRC64;

Query Match 38.5%; Score 1160.5; DB 2; Length 1171;
Best Local Similarity 42.9%; Pred. No. 4e-37;
Matches 271; Conservative 99; Mismatches 182; Indels 79; Gaps 20;

QY 1 SSDEEKND---VINNQSIINTDDNQIIKKEETN-----NYDGIKRSDETESTTNVD 51
DB 57 STENAKQDEASADNKVEVSETENNSTQKNDLTNPICKETNTDHSQEAAPTSTSTQQ 116
QY 52 ENEATPLQTPQDNTLHTEEEVKESSVSSNSIDTAQPSHTTINREESVQTSND 108
DB 117 QNNATTSTETQPN--IEKENVKPSTDKTATEDTSVILEEKKAPNNT--NNDVTTKPS-- 170
QY 109 VEDSHVSDFPANSKIKESNT---ESGKEENTIEQPNKVED---STTSQPSGYTNI--DEK 160
DB 171 -----TSEIQTPPTTPOESTNIENSQOPTSKVDNQVDTDTNPKPEPVNSKEEL 220
QY 161 ISNQDELLNLPINE--YENKARPLSTTSAQPSIKRV-----TVNQLAABQGSNNHLIK 212
DB 221 KNPKEKLKELVRNDSNTDRSTKPVATAPTAPSVAPKRVNAKIRFAVAQPAVASNNVDLIT 280
QY 213 VTDQSTIEGVDSEGVKAHDAENLIYDVTFEVDKVKSGDGTMTVDIDKNTVPSDLTDSF 272
DB 281 VTKQMITEGIKD-DGVIQAHGDEHIIYTSDFKIDNAVAKAGDTMTVKYDKETIPSDITDDF 339
QY 273 TTPKIKDNGSEIATGTVDNKNKQIITYTFTDYYVDKYENIKAHKLTSYIDSKVPNNNTK 332
DB 340 TPVDITDPSGEVJAKGTFDLNKTIIYKFTDYVDYRYENNAKLELSYIDKKEVP-NEIN 398
QY 333 LDVEYKTALSSVKNKTIITVEYQRPNEKRTANLQSMFTNIDTKNHTVEQTIYINPLAYSARE 392
DB 399 LNLTFATADKETSNNKVEYQKPIVKDESNIQSIIFSHLDTTKHEVEQTIYVNVPLKLNKN 458
QY 393 TNVNI-----SGNGD-----EGSTIIDSTIIKYKVGDNQNLPSNRIYDYSEYEDVTND 443
DB 459 TNVTIKSGGVADNGDYITGDSGTIIDSTNTEIKYKVASGQQLPQSNKIYDYSQYEDVTNS 518
QY 444 DYLAQLGNN--NDVNINFGNIDSPYIIKIVLSKYDPNKDDYTTIQQVTMTTINEYTGFE 500
DB 519 --VTINKNYGTNMANINFGDIDSAIYKVVSKVKTTPGAEDDLAVQQGVMTT-----TNKY 571
QY 501 RTASYDN-----TIAFSTSSGGQGGDLPPKTYKIGDYVWEDVDKDGIONTNDNEK 551
DB 572 NYSSYAGYITLFYQLLTIVVTVSVK-----PEEKLYKIGDYVWEDVDKVGQVQGTDSKEK 626
QY 552 PLSNVLVTIYDPDGTSKSVRTDEGDKYQFDG 582
DB 627 PMANVLVTIYDPDGTSKSVRTDANGHYEFGG 657

RESULT 5
```

```
086489
AC O86489 PRELIMINARY; PRT; 1166 AA.
ID
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sdr E protein.
GN SDR E.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Newman;
RX MEDLINE=99098700; PubMed=9884231;
RA Josefsson E., McCrea K., Ni Bihdin D., O'Connell D., Cox J., Hook M.,
RA Foster T.J.;
RT "Three new members of the serine-aspartate repeat protein multigene
family of Staphylococcus aureus."
RL Microbiology 144:3387-3395(1998).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY
CC AN AMIDE BOND (BY SIMILARITY).
DR EMBL; AJ005647; CAA06652.1; -.
DR PIR; T28680.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR008454; Cna_B.
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF05738; Cna_B; 3.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; YsIRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG anchor; 1.
DR TIGRFAMs; TIGR01168; YsIRK_signal; 1.
DR PROSITE; PS0847; GRAM_POS_ANCHORING; 1.
KW Cell wall; peptidoglycan-anchor.
SQ SEQUENCE 1166 AA; 126548 MW; 750A7B0135287D4A CRC64;

Query Match 38.0%; Score 1144.5; DB 2; Length 1166;
Best Local Similarity 42.6%; Pred. No. 1.6e-36;
Matches 263; Conservative 108; Mismatches 191; Indels 55; Gaps 18;

QY 1 SSDEKNDVINNQSIINTDNNQI--IKKETNNYDGIKESDRSTSTNVNDEATFL 58
DB 66 ATTSNKEVSEVENSTENNSTNNPIKGE--TNTDSQPEAKESSTSSQKQNNVTAT 123
QY 59 QKTPQDNTHLTEREVKES---SSVESNSSIDTAQPSHTTINKEESVQTSNDVDSHVS 115
DB 124 TEIKPON--IEKENVKSTDKTATEDSVILEEKKAPNT--NNDVTTPS-----TS 172
QY 116 DFANSKIKESNT---ESGKEENTIEQPNKVEDSTTSQPSGYTNIDEKIS--NODELLNLP 171
DB 173 EPSTSEIQTKPTTPQESTNIENSQPQTPSKVD---NQVTDATNPKEPVNVSKELKNPK 229
QY 172 -----INEYENKARPLSTTSQAQSIKRV-----TVNQLAAGQSNVNHILKVTQD 216
DB 230 EKULKELVRNSDNTDHSKTPVATAPTSVAPKRVNAKMFVAQAAPAAVANNVNDLIKVTQ 289
QY 217 SITEGYDDSEGVIAKHAENLIYDVTFEVDKVKSGDTMTVDIDKNTVPSDLTSDTTPK 276
DB 290 TIKVG-DGKNVAHAAGDKDIEYDTEFTIDNKVKGDTMTINTYDKNVPSDLTKNDPID 348
QY 277 IKNSGHIATGYDKNKQKQITFTFTDYVDKYNIAKHLKLTYSIDKSKVPNNNTKLDVE 336
DB 349 ITDPGSEVIAKGTDFDKATKQITFTFTDYVDKDYEDIKSRLTLYSIDKKTVP-NETSLNT 407
QY 337 YKTALSSVNKTIIVEYORPENENTANLOSMTFTIDTKNHVEQTIYINPLRYSAKENTYN 396
DB 408 FATAGKETSQNTVVDQDPMVHGDSNIQSIFTKLDEKQIEQOIYVNPPLKKSATNTKVD 467
QY 397 ISGNG-----DEGSTIIDSTIIKVKYKVDGNQMLPDSNRIYDYSEYEDVTND-DYA 446
DB 468 IAGSQVDYGNIKLNGSTIIDQNTTEIKVKYKVDQQLPQSNRIYDFSYEDVTSQFDNK 527

447 QLGNNNDVINFGNIDSPYIIKISKYDPNKDYTTIQOVTVMQTTINEYTGERTASYD 506
DB 528 KSFNNVATLDFGDIINSAYIIKVKYTPSTDSDELDAQCTSMRTT-DKY-GYNTAGYS 585
QY 507 NTIAFTSTSGQGQGLD-PPEKTYKIGDYVWEDVDKDGIGQNTNDNERKPLSNVLTLTYPDG 565
DB 586 NFIVTSNDTGGDGTVPKEKLYKIGDYVWEDVDKGVQGTDSKEKPMANVLTLTYPDG 645
QY 566 TSKSVRTDEGKYQFDG 582
DB 646 TTKSVRTDANGHYEFG 662

RESULT 6
Q8NXX5 PRELIMINARY; PRT; 1141 AA.
ID Q8NXX5
AC Q8NXX5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ser-Asp rich fibrinogen-binding bone sialoprotein-binding
DE protein.
GN SDR OR MW0518.
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA."
RL Lancet 359:1819-1827(2002).
DR EMBL; AF004823; BAB94383.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR008454; Cna_B.
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF05738; Cna_B; 3.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; YsIRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG anchor; 1.
DR TIGRFAMs; TIGR01168; YsIRK_signal; 1.
DR PROSITE; PS0847; GRAM_POS_ANCHORING; 1.
KW Complete proteome.
SQ SEQUENCE 1141 AA; 123997 MW; 372B5860850A332C CRC64;

Query Match 38.0%; Score 1143.5; DB 16; Length 1141;
Best Local Similarity 43.3%; Pred. No. 1.8e-36;
Matches 266; Conservative 98; Mismatches 192; Indels 59; Gaps 18;

QY 1 SSDEKNDVINNQSIINTDNNQI--IKKETNNYDGIKESDRSTSTNVNDEATFLQK 60
DB 69 SDNKEVSEVAENNSTENDSTNPI--KKETNT--DSQPEAKESSTKSSTQOQNNVTATTE 125
QY 61 TPQDNTHLTEREVKES---SSVESNSSIDTAQPSHTTINKEESVQTSNDVDSHVSDF 117
DB 126 TIKPON--IEKENVKSTDKTATEDSVILEEKKAPNT--NNDVTTPS----- 170
QY 118 ANSKIKESNT---ESGKEENTIEQPNKVEDSTTSQPSGYTNIDEKIS--NODELLNLP-- 171
DB 171 -TSEIQTKPTTPQESTNIENSQPQTPSKVD---NQVTDATNPKEPVNVSKELKNPK 226
QY 172 -----INEYENKARPLSTTSQAQSIKRV-----TVNQLAAGQSNVNHILKVTQD 218
DB 227 LKELVRNSDNTDHSKTPVATAPTSVAPKRVNAKMFVAQAAPAAVANNVNDLIKVTQ 286
QY 219 TEGYDDSEGVIAKHAENLIYDVTFEVDKVKSGDTMTVDIDKNTVPSDLTSDTTPK 278
DB 287 KVG-DGKNVAHAAGDKDIEYDTEFTIDNKVKGDTMTINTYDKNVPSDLTKNDPID 345
```



```

DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF05738; Cna_B; 3.
DR Pfam; PF04650; Gram_pos_anchor; 1.
DR Pfam; PF04650; YSIRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; YSIRK_signal; 1.
DR PROSITE; PS0847; GRAM_POS_ANCHORING; 1.
KW Complete proteome.
SQ SEQUENCE 1141 AA; 124038 MW; E679F7C2991846D9 CRC64;

Query Match
Best Local Similarity 37.8%; Score 1137.5; DB 16; Length 1141;
Matches 266; Conservative 97; Mismatches 193; Indels 59; Gaps 19;

QY 1 SSDEKNDVNNQSIINTDNNQIIKKEETNNYDGIKESDRTESTTNVDENEATFLQK 60
DB 69 SDNKEVSVETNNSTTENDSNPI--KKEIWT--DSQPEAKEESTTSQQQNNVTATTE 125
QY 61 TPQNTHLTHEREVKES---SSVESNSIDTAQOPSHHTINREESVQTSNVEDSHVSD 117
DB 126 TKPQN--IERENVKPSDKTATEDTSVLEKAPNYT--NNDVTTKPS----- 170
QY 118 ANSRIKESNT---ESGKEENTIEQPNKVEDSTTSQPSGYTNIDKIS--NODELLNLP-- 171
DB 171 -TSRIQKPTTPQESTNIENSQOPTESKVD---NQVTDATNPKPEVNVSKLELNKPEK 226
QY 172 -----INEYENKARPLST--TSAQP-----SIKRVTVNQLAAEQGSNVNHLIKVTDQSI 218
DB 227 LKELVRDNNNTDRSTKPVATAPTSVAPKRLNAKMPFAVAPAAVASNNVNDLITVTKQTI 286
QY 219 TEGYDDSEGVKKAHAENLIYDVTFFVDDKVKSGDWTVDIDKNTVPSDLTDSFTIPKIK 278
DB 287 KVG-DGKDNVAAAHGDKDIEYDTEFTIDNKVKSGDWTMTINYNKKNVIPSDLTDRKNDPIDIT 345
QY 279 DNSGEIATGYDNKNKQIITFTDYVDKYENIKAKHLKLTYSIDKSKVPNNNTKLDVEYK 338
DB 346 DPSEVIAGTFDKATQIITFTDYVDKYEDIKARLTLSYIDKQAVP-NEISLNTFA 404
QY 339 TALSSVNTKITVEYQRPENRTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVNIS 398
DB 405 TAGKETSONSVQYDPMVHGDSNIQSIFTKLDENKQTIIEQIYVNPFLKKTATNTKVDIA 464
QY 399 GNG-----DEGSTIIDSIIKVKYKQDNQNLPSNRIDYSEYEDVTND-DYAQL 448
DB 465 GSQVDDYGNIKLNGSGTIIDONTEIKYKYNPNQQLPQSNRIYDFSQYEDVTSQFDNKK 524
QY 449 GNNNDVNFNGNIDSPYIIKVISKYPDNKDDYTTIQQVTVMQTTINBYTGEFTASVDNT 508
DB 525 FSNNAVLDFGDNINSALIIKVSQKPTSDGELDIAGTSMRTT-DKY-GYNNVAGSYNF 582
QY 509 IAFSTSSGQGGDL-PPEKTYKIGDYVWEDVDKGIQNTNDNEKPLSNVLVTLTYPDGTS 567
DB 583 IVTSNDTGGGDTGVKPEKLYKIGDYVWEDVDKGVQGTDSKEKPMANVLVTLTYPDGTT 642
QY 568 KSVRTDEGKYQPDG 582
DB 643 KSVRTDANGHYEFG 657

RESULT 9
O86488 PRELIMINARY; PRT; 1315 AA.
AC O86488;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Srd protein.
GN SRD.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=Newman;
RX MEDLINE=9908700; PubMed=9884231;
RA Josepson E., McCrea K., Ni Eidhin D., O'Connell D., Cox J., Hook M.,
RA Foster T.J.;
RT "three new members of the serine-aspartate repeat protein multigene
RT family of Staphylococcus aureus.";
RL Microbiology 144:3387-3395(1998).
CC -|- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY
CC AN AMIDE BOND (BY SIMILARITY).
DR EMBL; AJ005646; CAA06651.1; -.
DR PIR; T28679; T28679.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR008454; Cna_B.
DR InterPro; IPR005877; Gpos_YSIRK.
DR Pfam; PF05738; Cna_B; 5.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; YSIRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; YSIRK_signal; 1.
DR PROSITE; PS0847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 1315 AA; 142775 MW; 58D4E1F48EE6A689 CRC64;

Query Match
Best Local Similarity 30.1%; Score 576; DB 2; Length 1315;
Matches 186; Conservative 112; Mismatches 228; Indels 92; Gaps 29;

QY 1 SSDEKNDV---INNQSINTDNNQIIKKEETNNYDGIKESDRTESTTNVDE-NEAT 56
DB 55 STNKELNEATTSASDNQSSDKVDMQOLNQEDNTKNDNQKEMVSSQGNETTSNGNKLIEKE 114
QY 57 FLQKTPQDNTHLTHEREVKESSSVESNSIDTAQOPSHHTINREESVQTSNVEDSHVSD 116
DB 115 SVQSTTGKVFVSTAKSDEQASPKSTNEDLNTKQ-----TISNQELQ-PDIQENKSVVN 168
QY 117 FANSKIKESNTSGKEENTIEQPNKVEDSTTSQPSGYTNI-DEKISNODELL--NLPI 173
DB 169 -----VQPTNEENKQVD-----AKTESTT-----LNKSDAIKSNDELTVDNNSNS 209
QY 174 EYENKARPLSTSAQP-----SIKRVTVNQLAAEQGSNVNHLIKVTDQSIPEGYDDSEGV 229
DB 210 NNENNADIILPKSTAPKRLNTRMRIAQVPSSTEAQNVNDLITSNTTLTVVDADKNNKIV 269
QY 230 KAHAENLIYDVTFFVDDKVKSGDWTVDIDKNTV-----PSDLTDSFTIPKID-NSG 282
DB 270 PAQDYLILSKQIT--VDDKVKSGDYFTIKY-SDTVQVYGLNPFEDIKN---IGDIKDPNNG 323
QY 283 EIIATGYDNKNKQIITFTDYVDKYENIKAKHLKLTYSIDKSKVPNNNTKLDVEYKTALS 342
DB 324 ETIATAKHDTANNLITYTFTDYVDRFNSVQMGINSIYMDADTIP--VSKNDVEFNVTIG 381
QY 343 SVNKITITVEYQRPN--ENRTANLQSMFTNIDTKNHT-----VEQTIYINPLRYSAKE 392
DB 382 NTTTKTTANIQYDPYVNVNKNKSGSAFT--ETVSHVGNKENFGYKQTIYVNPSENLTN 439
QY 393 TNVNI-----SGNGDEGSTIIDSIIKVKYKQDNQNLPSNRIDYVS--EYEDVTNDVYA 446
DB 440 AKLKQVAYHSSYPNNGIQINKDVTDIKIYQVPGYTL---NKGVDNTEKELTDVTN-QYL 495
QY 447 Q---LGNNDVNFNGNIDSPYIIKVISKYPDNKDDYTTIQQVTVMQTTINBYTGEFTFA 503
DB 496 QKITTYGDNNSAVIDFGNADSAVVMVNTKFTQNTSESPTLVQWATLSSGTN-----KSV 549
QY 504 SYDNTIAFSTSSGQGGDL-PPEKTYKIGDYVWEDVDKGIQNTNDNEKPLSNVLVTLTY 563
DB 550 STGNALGFTNNQSGAG-----QEVYKIGNYVWEDTNKGVQEL--GEKGVNVTVTV-FD 602
QY 564 DGTSKSVR---TDEDGKY 578
DB 603 NNTNTKVGCAVTKEDGSY 620

```



QY 176 ENKARPLSTTGAQPSIKKVTNQAAEQGS-----NVNHLIKVTDQSTIGTEGVDDSEGVIKA 231  
 DB 212 ENNADIILPKSTAPKSLNTRMAAIOQNSDTSKXNVNDLITNTTLTVVDADNSKTIKTPA 271  
 QY 232 HDAENLIYDVTFEVDKVKSGDVTVDIDKNTV-----PSDLTDSFTIPKID-NSGEI 284  
 DB 272 QDYLKSKQIT--VDDKVKSGDYFTIKY-SDTQVYGLNPEDIKN---IGDIKDPNNGET 325  
 QY 285 IATGYDNKKNQIYTFDYDYDKYENIKAHKLKLSYIDKSKVPNNNTKLDVVEYKLTALSSV 344  
 DB 326 IATAKHDTANNLIYTFDYDYDRNSVKMGINSYIYMDADTIPVD--KQVFPFSVTIGNQ 383  
 QY 345 NKTIYVEYQRP--NENRTANLQSMFTNIDTKNHT-----VQTYIINPLRYSAKETN 394  
 DB 384 ITTTTADITYPAYKEADNNSGSAFT--ETVSHGVNVEDPGYVQVYVVPNMDKDLKGAK 441  
 QY 395 VNISNGDEGST---IIDDSTILKVKYKVDNQNLDPDSNRIDY--SEYEDVND--DYA 446  
 DB 442 LKVEAYHPKYPTNIGQINQNVNLIKIRVPEGYTL---NKGVDVNTDLVDVDFEKNKM 498  
 QY 447 QLGNNNDVNINFGNIDSPYIKIVISKYDPNKKDDYTTIQQVTMTTINEXYTGFEFRASVD 506  
 DB 499 TYGSNQSVNLDFDITSAIVVMVNTKQYTNSESPTLVQMATLSGK-----KSVSTG 552  
 QY 507 NTIAFSTSSGGQGLDPEKTYKIGDYVWEDVDKQIGQNTNDNEKPLSNVLVLTYPDGT 566  
 DB 553 NALGFTNNQSGAG---QEVYKIGNYVWEDTNKNGVQEL--GEKGVGNVTIV--FDNNT 605  
 QY 567 SKSVR---TDEGKY 578  
 DB 606 NTKVGEAVTKEDGSY 620

## RESULT 12

Q9K114 ID Q9K114 PRELIMINARY; PRT; 1733 AA.  
 AC Q9K114;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative cell-surface adhesin SdrF.  
 GN SDRF.  
 OS Staphylococcus epidermidis.  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OX NCBI\_TaxID=1282;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=9491;  
 RX MEDLINE=20340957; PubMed=10878118;  
 RA McCrea K.W., Hartford O., Davis S., Ni Eidhin D., Lina G.,  
 RA Speziale P., Foster T.J., Hook M.;  
 RT "The serine-aspartate repeat (Sdr) protein family in Staphylococcus  
 epidermidis."  
 RL Microbiology 146:1535-1546 (2000).  
 DR EMBL: AF245041; AAF72509.1; --  
 DR GO: GO:0016020; C:membrane; IEA.  
 DR InterPro: IPR008454; Cna\_B.  
 DR InterPro: IPR005877; Gpos\_Ysirk.  
 DR InterPro: IPR001899; Gram\_pos\_anchor.  
 DR Pfam: PF05738; Cna\_B; 4.  
 DR Pfam: PF00746; Gram\_pos\_anchor; 1.  
 DR Pfam: PF04650; Ysirk\_signal; 1.  
 DR PROSITE: PS00847; GRAM\_POS\_ANCHORING; 1.  
 SQ SEQUENCE 1733 AA; 184720 MW; DBD62EA1692FD4E8 CRC64;

Query Match 16.6%; Score 499; DB 2; Length 1733;  
 Best Local Similarity 27.1%; Pred. No. 1.5e-11;  
 Matches 168; Conservative 104; Mismatches 275; Indels 72; Gaps 25;  
 QY 1 SDEEKNDVNNQSIINTDNNQIILKEEFNNYDGIKESDRTSTTNDVNEA-TFLQ 59  
 DB 148 TSTTQDDSTKKNPSLKNLNLSSSTTSKESKTDHSTKQAMST-NKSNLDITNDSPTQSE 206

QY 60 KT-PQDNTHLTEREVKSSSVSSNS-----SIDTAQQPSTHTIN-REESVOTSDN 108  
 DB 207 KTSQANNDSTDNQSAPSKQLDSKPSQKVKYKTKFENDEPTQDVEHTTKLKTFSVSDSS 266  
 QY 109 VESHVSDDFPANSKIKESNTESGKEENYIROPNKKVSDSTTSQSPGYNTIDDKI----- 161  
 DB 267 VNDK--QDYTRSAV-----ASLGVDSNETEAITNAVRDNLDLKAASREQINEAIAEALKK 320  
 QY 162 --SNOBELNLNP--INEYENKARPLSTTSAQPSIKKVTNQAAE--QGSNVNHLIKVTD 215  
 DB 321 DFNPDYGVDTPLALNRSQSKSP--HKASP---RNLMSLAEPNSGKNVNDKVKITN 375  
 QY 216 Q--SITEGVDDSGVKAHDAENLIYDVTFEVDKVKSGDVTVDIDKNTVPSDLTDSFT 273  
 DB 376 PTLNLKSNHANNHNVWPTSNQFNLKANYELDDSIKEGDTFTIKYQYIRPGGLELPAI 435  
 QY 274 IPKIKNSGHEIATGYDNKKNKOITTTFTDYVDKYENIKAHKLKLSYIDKSKVPNNNTKL 333  
 DB 436 KTLRSKDSISVANGVYDKTNTTNTTFTYTYVDQYQNTITGSDFLIATPKRETAIKQNY 495  
 QY 334 DVEYKALSGVNKTIIVYQRPNNENRTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKET 393  
 DB 496 PHEVTIANEVVKDFIVDGNKKONTT---AAVANVDNVNKHNEVVLNQNQPKYA 552  
 QY 394 NVNISNGDEGSTIIDSTI---IKYKVGDMQNLPS-NRIYDYSEYEDVNDYDQAL- 448  
 DB 553 KY-----FSTVKNGEFIPGEVKYVEVTDNAMVDSFNPDLNNSNVKDVTSQFAPKVS 604  
 QY 449 GNNNDVNINFGNI---DSPYIKIVISKYDPNKKDDYTTIQQVTMTTINEXYTGFEFRASV 505  
 DB 605 ADGTRVDINPASMANGKYYIVTQAVRPTGTGNVYT--EYWLTRDGTN--TNDFYRGTK 660  
 QY 506 DNTIAFSTSSGGQGLDPEKTYKIGDYVWEDVDKQIGQNTNDNEKPLSNVLVLTYPDG 565  
 DB 661 SITVTYLGSSSTAQGNP---TVSLGYVWLDKKNQVQ--DDDEKGLAGVYVTL--KDS 713  
 QY 566 TSKS---VTDDEGKYQFD 581  
 DB 714 NNRELQRTVTDQSGHYQFD 732

RESULT 13  
 Q8CMP4 ID Q8CMP4 PRELIMINARY; PRT; 1633 AA.  
 AC Q8CMP4;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding  
 DE protein.  
 GN SE2395.  
 OS Staphylococcus epidermidis.  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OX NCBI\_TaxID=1282;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 12228;  
 RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,  
 RA Chen Z., Wen Y.;  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AS016752; AAO06038.1; --  
 DR GO: GO:0016020; C:membrane; IEA.  
 DR InterPro: IPR008454; Cna\_B.  
 DR InterPro: IPR005877; Gpos\_Ysirk.  
 DR InterPro: IPR001899; Gram\_pos\_anchor.  
 DR Pfam: PF05738; Cna\_B; 4.  
 DR Pfam: PF00746; Gram\_pos\_anchor; 1.  
 DR Pfam: PF04650; Ysirk\_signal; 1.  
 DR TIGRFAMs: TIGR01167; LPXTG\_anchor; 1.  
 DR TIGRFAMs: TIGR01168; Ysirk\_signal; 1.  
 DR PROSITE: PS00847; GRAM\_POS\_ANCHORING; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 1633 AA; 174768 MW; DAA8537B8CD288BD CRC64;

```
Query Match 16.5%; Score 497; DB 16; Length 1633;
Best Local Similarity 27.0%; Pred. No. 1.7e-11;
Matches 167; Conservative 105; Mismatches 275; Indels 72; Gaps 25;

QY 1 SSDEKNDVNNQNSINTDNNQI IKKEETNNYDGIKESDETESTTNDENEATFLQ 59
DB 148 TSTQDSTKSNPSLKDNLNSSTTSKESKTDHSTKQAMST-NKSNLDNDSTQSE 206

QY 60 KT-PQDNTLHTEBEVKESSSVSSNS-----SIDTAQPSHTTIN-REESVOTSN 108
DB 207 KTSQANDSTNQSPSKQSDKSPSEKQVYKTKENDEPTQDVHETTKLTKTSISTDSS 266

QY 109 VEDSHVSDPANSIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYNTIDEKI----- 161
DB 267 VNDK--QDYTRSAV---ASLGVDNSNETAITNAVRDNLDLKAASREQINEAIIAELKK 320

QY 162 --SNQDELLNLP--INEYENKARPLSTSAQPSIKKVTYNQLAAE--QGSNNHLIKVTD 215
DB 321 DFENPDYGVDTPLALNTSQSKNP--HKSASP---RNNLSLAAEPNSGKNVNDKVKITN 375

QY 216 Q--SITEGVDSSEGVKAHDAENLIYDVTPEVDKVKSGDTMTVDIDKNTVPSDLTDSFT 273
DB 376 PTLNLSNNHANNVIWPTSNQFNKANYELDDSIKEDTFTIKYGYIRPGLELPAI 435

QY 274 IPKIDNSGELIATGYDNNKQKQITFTFTDYDVKYENIKAHKLKLTSYIDSKVPNNNTKL 333
DB 436 KTLRSKQSGISVANGVYDKTNTTFTFTYDQYQNTGSGFDLIATPKRETAIKDNQNY 495

QY 334 DVEYKTLSSVNTITVEQRPNEENTANLQSFNTIDTKHTVEQTIYINPLRYSAKET 393
DB 496 PMEVTIANEVKKDFIVDYGKNDKNTT---AAVANVDNNKHNEVYLNQNNQPKYA 552

QY 394 NVNISGDEGSTIIDSTI---IKVKYKGDQNLPS-NRIYDYSEYEDVTDNDYAOQL- 448
DB 553 KY-----FSIVKNGKFIPEGVKYEVYDITNAWVDSFNPDLNSSNVKDTISQFTPKYS 604

QY 449 GNNNDVNIINPNI---DSPYIIKISKYDNPDKDDYTIQQTVMQTTINEYGEFTASY 505
DB 605 ADGTRVDIINPARSWANGKYIVQAVRPTGNGVYT--EYWLTRDGTN--TNDYRGTK 660

QY 506 DNTIATSSGQGGDLPEPKTYKIGDYWEDVDKGIQNTDNEKPLSNVLTLYTPQG 565
DB 661 STVTYLSGSSTAQGNP---TSLGDBYVWLDKNGVQ--DDDEKGLAGVYVTL--KDS 713

QY 566 TSKS---VRTEDEKTYQFD 581
DB 714 NNRELQRTVTDQSGHYQFD 732

RESULT 14
Q99W48 PRELIMINARY; PRT; 953 AA.
AC Q99W48
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding
DE protein.
GN SDRS OR SAV0561 OR SA0519.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=1418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
```

```
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.",
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003359; BAB56723.1; -.
DR EMBL; AP003131; BAB41750.1; -.
DR PIR; C89824; C89824.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_YSIK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF05738; Cna_B; 2.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; YSIK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; YSIK_signal; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Complete proteome.
SQ SEQUENCE 953 AA; 103292 MW; 729A7169A074A1E5 CRC64;

Query Match 16.2%; Score 489; DB 16; Length 953;
Best Local Similarity 28.2%; Pred. No. 2e-11;
Matches 162; Conservative 82; Mismatches 236; Indels 94; Gaps 22;

QY 35 GIEKSEDETESTTNDENEATFLQPTDNTLHTEBEVKESSSVSSSIDTAQPSH 94
DB 45 GHEAKAAHTGELNQSQKNETT---APSENK--TTEKV--DSRQLKNTQTATADQPKV 96

QY 95 TTINREESVQTSNDVEDSHVSDPANSIKESNTESGKEENTIEQPNKVKEDSTTSQSGY 154
DB 97 T-----MSDSATVKTSTSNQ-----PONATASQSTTQTSNV 129

QY 155 TNIDEK---ISNQDELLNLPINEYEN-KARPLSTTSAQPSIKKVTYNQLAA-EQGSNVNH 209
DB 130 TTNDKSSITYSNETDKSNL--TQAKNVSTPTKTTIKQALANRMAVNTVAAPQGINVND 187

QY 210 LIKVTDDQSI-----TEGYDDESEGVKAHDAENLIYDVTFEVDDKVKSGDTMT 256
DB 188 KVHFTNIDIAIDKGHVNTKTGNTFEWATSSDLK-----LKANYTIDDSYKEGDTFT 239

QY 257 VIDKNTVPSDLTDSFTIPKIDNSGELIATGYDNNKQITFTFTDYDVKYENIKAHKL 316
DB 240 FKGQYFREGSVRLPSQTLNQAQNIIAKGYDSKNTTFTFTYDQYTNVSGSFE 299

QY 317 LTSYIDSKSVNNNTKLDVEYKTLSSVNNKTTTVEYQRPNEENTANLQSFNTIDTKNHT 376
DB 300 QVAFAKRENATTDKTAKEVTLGNDTYSKDIVDY--GNQKGOQLISTNVINNEEDLS 356

QY 377 VEQTIYIN-PLRYSAKETNV-NISGNGDEGSTIIDSTIIKVKYKGDQNLPSNRIYDY 434
DB 357 RNMVYVNPQKTYTKETFTVNTL-----GYKFNPDANKFKIYEVDQNFVDS-FTPDT 410

QY 435 SEYEDVTND-DYAOQLGNNNDVNIENFN-----TDSPIIKVKISKYDPNKDDYTIQQTVM 489
DB 411 SKLDVGTQGFVYNDNKNKTATVDLLNGSSDKQYIIQOAVPDPNSSTDNGKIDYTLT 470

QY 490 QTTINEYTGFEFTASYDNTIATFSTSGQGGDLPEPKTYKIGDYWEDVDKGIQNTDND 549
DB 471 QNGKSSWSN-----SYSNVNGSSTANGD-----QKKYNLGDYWEDINKGQKDA--N 516

QY 550 EKPLSNVLTLYTPDQTS-KSVRTDDEGKYQFDG 582
DB 517 EKGIGKVYVILKDSNGKELDRITTTDENGKYQFTG 550

RESULT 15
Q8KWM1 PRELIMINARY; PRT; 1893 AA.
ID Q8KWM1
AC Q8KWM1
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:33:49 ; Search time 122.638 Seconds  
(without alignments)  
2515.875 Million cell updates/sec

Title: US-10-806-288-15  
Perfect score: 5645  
Sequence: 1 MINKNNLLTKKKPIANKSN.....FAGLGALLGKKRKKNKN 1092

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003Bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5646	100.0	1092	2 AAW41602	Aaw41602 Staphyloc
2	5646	100.0	1092	7 ABM79019	Abm79019 Staphyloc
3	4490	79.5	930	5 ABP40469	Abp40469 Staphyloc
4	4485	79.4	991	3 AAY83171	Aay83171 Cell wall
5	4485	79.4	991	3 AAY70120	Aay70120 Staph. ep
6	4307	76.3	892	6 ABU42557	Abu42557 Protein e
7	3351.5	59.4	670	6 ABU42520	Abu42520 Protein e
8	2787	49.4	549	7 ABM79020	Abm79020 Staphyloc
9	2698.5	47.8	1166	2 AAY08643	Aay08643 S. aureus
10	2698.5	47.8	1166	6 ABJ18982	Abj18982 Pathogen
11	2624.5	46.5	560	7 ABM79015	Abm79015 Staphyloc
12	2586	45.8	1141	6 ABU42327	Abu42327 Protein e
13	2339.5	41.4	1633	6 ABU42513	Abu42513 Protein e
14	2303.5	40.8	1802	3 AAY83170	Aay83170 Cell wall
15	2303.5	40.8	1802	3 AAY70119	Aay70119 Staph. ep
16	2256.5	40.0	995	6 ABM72437	Abm72437 Staphyloc
17	2234	39.6	1385	6 ABU16400	Abu16400 Protein e
18	2205.5	39.1	1920	6 ABU43489	Abu43489 Protein e
19	2093.5	37.1	953	6 ABU16533	Abu16533 Protein e
20	2070	36.7	1349	6 ABM72436	Abm72436 Staphyloc
21	2069	36.6	1349	4 AAU37544	Aau37544 Staphyloc
22	2069	36.6	1349	4 AAU34402	Aau34402 Staphyloc
23	2032.5	36.0	947	6 ABJ18940	Abj18940 Pathogen
24	2011	35.6	932	4 AAU36845	Aau36845 Staphyloc
25	2011	35.6	932	4 AAU34082	Aau34082 Staphyloc

26	1990	35.2	930	2 AAY08641	Aay08641 S. aureus
27	1880	33.3	1315	2 AAY08642	Aay08642 S. aureus
28	1880	33.3	1315	6 ABJ18969	Abj18969 Pathogen
29	1672	29.6	343	7 ABM79016	Abm79016 Staphyloc
30	1634	28.9	1021	4 AAU33975	Aau33975 Staphyloc
31	1634	28.9	1021	4 AAU36951	Aau36951 Staphyloc
32	1631	28.9	933	6 ABJ18947	Abj18947 Pathogen
33	1629.5	28.9	936	2 AAW89801	Aaw89801 Staphyloc
34	1628	28.8	933	3 AAY58435	Aay58435 Staphyloc
35	1628	28.8	933	4 AAB69508	Aab69508 Staphyloc
36	1625	28.8	927	6 ABM72221	Abm72221 Staphyloc
37	1617.5	28.6	918	2 AAY08640	Aay08640 S. aureus
38	1611	28.5	913	6 ABJ18917	Abj18917 Pathogen
39	1581	28.0	935	6 ABU16402	Abu16402 Protein e
40	1562	27.7	316	7 ABM79017	Abm79017 Staphyloc
41	1552	27.5	877	6 ADA89539	Ada89539 Staphyloc
42	1552	27.5	877	6 ABM72702	Abm72702 Staphyloc
43	1536	27.2	877	6 ABU42504	Abu42504 Protein e
44	1446	25.6	287	4 AAG81687	Aag81687 S. epider
45	1441.5	25.5	841	4 AAU37158	Aau37158 Staphyloc

ALIGNMENTS

RESULT 1  
AAW41602  
ID AAW41602 standard; protein; 1092 AA.  
XX  
AC AAW41602;  
XX  
DT 17-OCT-2003 (revised)  
DT 22-JUN-1998 (first entry)  
XX  
DE Staphylococcus epidermidis fibrinogen binding protein FIG.  
XX  
KW Fibrinogen binding protein; FIG; aggregation; infection;  
KW coagulase-negative Staphylococcus; therapy; diagnosis; immunisation;  
KW immunogen; vaccine.  
XX  
OS Staphylococcus epidermidis; strain HB.  
XX  
PH Key Location/Qualifiers  
FT Peptide 1..51  
FT Protein 52..1092 /label= Sig\_peptide  
FT Region 52..824 /label= Mat\_protein  
FT /note= "non-repetitive region, harbours fibrinogen binding activity"  
FT Region 825..1040  
FT /note= "Asp-Ser dipeptide repeat region"  
FT Region 1053..1057  
FT /note= "cell wall anchoring motif"  
XX  
WO9748727-A1.  
XX  
24-DEC-1997.  
XX  
18-JUN-1997; 97WO-SE001091.  
XX  
20-JUN-1996; 96SE-00002496.  
XX  
(GUSS/) GUSS B.  
(NILS/) NILSSON M.  
(FRYK/) FRYKBERG L.  
(FLOC/) FLOCK J.  
(LIND/) LINDBERG M.  
XX  
Guss B, Nilsson M, Frykberg L, Flock J, Lindberg M;  
WPI; 1998-063079/06.  
DR N-PSDB; AAV04279.

XX Fibrinogen-binding protein from coagulase-negative Staphylococcus - used  
PT for prevention, treatment and diagnosis of Staphylococcus infection.

XX Example 3; Fig 6; 45pp; English.

XX The protein comprises the fibrinogen binding protein (FIG) of coagulase-  
CC negative Staphylococcus epidermidis HB. Its amino acid sequence was  
CC deduced from the isolated fig gene (see AAV04279). The closest known  
CC analogue of FIG is the clumping factor of *S. aureus* which also binds  
CC fibrinogen and promotes bacterial aggregation in serum. Recombinant FIG  
CC polypeptides can be expressed in host cells. They are used as immunogens,  
CC particularly in vaccines (which may be expressed in vivo) to protect  
CC humans and animals against coagulase-negative Staphylococcus infection.  
CC Antibodies raised against FIG can be used for passive immunisation. They  
CC block the adherence of bacteria) and for diagnosis. (Updated on 17-OCT-  
CC 2003 to standardise OS field)

XX Sequence 1092 AA;

Query Match 100.0%; Score 5646; DB 2; Length 1092;  
Best Local Similarity 100.0%; Pred. No. 1.3e-266;  
Matches 1092; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MINKNNLLTKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNKAKABENSVDVK 60  
DB 1 MINKNNLLTKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNKAKABENSVDVK 60  
QY 61 DSNITDDELSDNSQSSDEEKNDVINNNQSIINTDNNQI1KKEETNNYDGIKESDRTES 120  
DB 61 DSNITDDELSDNSQSSDEEKNDVINNNQSIINTDNNQI1KKEETNNYDGIKESDRTES 120  
QY 121 TTNVDENEATFLOKTPQDNTHLTHEEVEKSSSVSSSIDTAQPSHTTINREESVQTS 180  
DB 121 TTNVDENEATFLOKTPQDNTHLTHEEVEKSSSVSSSIDTAQPSHTTINREESVQTS 180  
QY 181 DNVEDSHVSPANKIKESNTESGKEENTIEQPNKVEDSTTQPSGYTNIDKISNQDE 240  
DB 181 DNVEDSHVSPANKIKESNTESGKEENTIEQPNKVEDSTTQPSGYTNIDKISNQDE 240  
QY 241 LLNLPINEYENKARPLSTTSAQPSIKRVTNQLAAEQGSNNVHLIKVTDQSIITEGYDDSE 300  
DB 241 LLNLPINEYENKARPLSTTSAQPSIKRVTNQLAAEQGSNNVHLIKVTDQSIITEGYDDSE 300  
QY 301 GV1KAHDAENLIYDVTPEVDKVKSGDTMTVDIDKNTVPSDLTDSFTPIKIDNSGEIIA 360  
DB 301 GV1KAHDAENLIYDVTPEVDKVKSGDTMTVDIDKNTVPSDLTDSFTPIKIDNSGEIIA 360  
QY 361 TGYDKNKQITFTDYVDKYENIKAHKLTSYIDKSKVPNNNTKLDVEYKTALESVVK 420  
DB 361 TGYDKNKQITFTDYVDKYENIKAHKLTSYIDKSKVPNNNTKLDVEYKTALESVVK 420  
QY 421 TITVEYORPNENRTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVIISNGDEGST 480  
DB 421 TITVEYORPNENRTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVIISNGDEGST 480  
QY 481 IIDSTIIKVKYKGDNQLPDSNRIYDYSEYEDVTNDDYAQLGNNDVNIIFNGNIDSPYI 540  
DB 481 IIDSTIIKVKYKGDNQLPDSNRIYDYSEYEDVTNDDYAQLGNNDVNIIFNGNIDSPYI 540  
QY 541 IKVISKYDPNKDDYTTTQOVTWQTINNEYTGFRASYNNTAFSTSSQGGQGLDLPPEK 600  
DB 541 IKVISKYDPNKDDYTTTQOVTWQTINNEYTGFRASYNNTAFSTSSQGGQGLDLPPEK 600  
QY 601 TYKIGDVWEDVDKDG1QNTNDNEKPLSNVLVTLTYPDGSKSVRTDDECKYQFDGLKNG 660  
DB 601 TYKIGDVWEDVDKDG1QNTNDNEKPLSNVLVTLTYPDGSKSVRTDDECKYQFDGLKNG 660  
QY 661 LTYKITPETPEGYTPTLKHSGTNPALDSEGNVSVWVTINGQDDMTIDSGFYQTPKYSIGNY 720  
DB 661 LTYKITPETPEGYTPTLKHSGTNPALDSEGNVSVWVTINGQDDMTIDSGFYQTPKYSIGNY 720  
QY 721 VWDTNKG1QGDDEKIGSVKVTLDKENGNI1STTTTDENGKYQFDNLNSGNYIVHFDK 780

DB 721 VWDTNKG1QGDDEKIGSVKVTLDKENGNI1STTTTDENGKYQFDNLNSGNYIVHFDK 780  
QY 781 PSQWTTQTTTSDGDDDEQDADGEEVHTITDHDDFSIDNGYDDDESDDSDSDSDSD 840  
DB 781 PSQWTTQTTTSDGDDDEQDADGEEVHTITDHDDFSIDNGYDDDESDDSDSDSDSD 840  
QY 841 SD 900  
DB 841 SD 900  
QY 901 SD 960  
DB 901 SD 960  
QY 961 SD 1020  
DB 961 SD 1020  
QY 1021 SD 1080  
DB 1021 SD 1080  
QY 1081 LGKRRKRNKNKN 1092  
DB 1081 LGKRRKRNKNKN 1092  
RESULT 2  
ABW79019  
ID ABW79019 standard; protein; 1092 AA.  
XX  
AC ABW79019;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Staphylococcus epidermidis polypeptide.  
XX  
KW Infection; antibacterial; vaccine.  
XX  
OS Staphylococcus epidermidis.  
PN WO2003076470-A1.  
XX  
PD 19-SEP-2003.  
XX  
PF 05-MAR-2003; 2003WO-US006415.  
XX  
PR 05-MAR-2002; 2002US-0361324P.  
XX  
PA (INHI-) INHIBITEX INC.  
PA (TEXA) UNIV TEXAS A & M SYSTEM.  
XX  
PI Patti JM, Hutchins JT, Hall A, Domanski P, Patel P, Hook M;  
PI Robbins J, Vernachio J, Bowden MG;  
XX  
DR WPI; 2003-722324/68.  
XX  
PT New antibody recognizing a Staphylococcus epidermidis protein comprising  
PT Sdrg N1X2N3, Sdrg N2N3 or Sdrg2 useful for preparing a composition for  
PT treating or preventing a coagulase-negative Staphylococcus infection.  
XX  
PS Claim 16; Page 36-37; 78pp; English.  
XX  
CC The present sequence comprises the protein sequence of a polypeptide of a  
CC coagulase-negative Staphylococcus epidermidis. A claimed monoclonal  
CC antibody recognises this protein and is used in a claimed method of  
CC treating or preventing a coagulase-negative staphylococcal infection in a  
CC human or animal, e.g. a nosocomial coagulase-negative staphylococcal  
CC infection in low birth weight infants  
XX  
SQ Sequence 1092 AA;

Query Match		100.0%; Score 5646; DB 7; Length 1092;
Best Local Similarity		100.0%; Pred. No. 1.3e-266;
Matches 1092; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	MINKNNLLTKKPIANKSNKYAIRFTVGTASIVIGATLLFGLGNEAKAENSVDVK 60
DB	1	MINKNNLLTKKPIANKSNKYAIRFTVGTASIVIGATLLFGLGNEAKAENSVDVK 60
QY	61	DSNTDDELSDNSDQSSDEEKNDVINNNQSIINTDDNNQIIKKEETNNYDGIKRSERDTS 120
DB	61	DSNTDDELSDNSDQSSDEEKNDVINNNQSIINTDDNNQIIKKEETNNYDGIKRSERDTS 120
QY	121	TTNVDENEATFLQKTPQDNTLHTEEEVKSSSVSSSSIDTAQPSHTTINREESVQTS 180
DB	121	TTNVDENEATFLQKTPQDNTLHTEEEVKSSSVSSSSIDTAQPSHTTINREESVQTS 180
QY	181	DNVEDSHVDFPANSKIKESNTSGKENTIEQPNKVKESTTSQPSGYTNIDKISNQDE 240
DB	181	DNVEDSHVDFPANSKIKESNTSGKENTIEQPNKVKESTTSQPSGYTNIDKISNQDE 240
QY	241	LLNLPINEYENKARPLSTTSAQPSIKRVTVNQLAAEQGSNNVHLIKVTDQSITEGYDDSE 300
DB	241	LLNLPINEYENKARPLSTTSAQPSIKRVTVNQLAAEQGSNNVHLIKVTDQSITEGYDDSE 300
QY	301	GVIAKAHAENLLIYDVTPEVDVKVSGDTMTVDIDKNTVPSDLTDSFTPIKIDNSGEIIA 360
DB	301	GVIAKAHAENLLIYDVTPEVDVKVSGDTMTVDIDKNTVPSDLTDSFTPIKIDNSGEIIA 360
QY	361	TGYDNNKQIITVFTDYDVKYENIKAHKLTSYIDKSKVPNNNTKLDVEYKTALESVVK 420
DB	361	TGYDNNKQIITVFTDYDVKYENIKAHKLTSYIDKSKVPNNNTKLDVEYKTALESVVK 420
QY	421	TITVEYQRPNEANTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVNISGDEGST 480
DB	421	TITVEYQRPNEANTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVNISGDEGST 480
QY	481	IIDDSITIIKVKYQDGNQLPDSRIYDYSEYEDVTHDDYAQLGNNDVNIINFGNIDSPYI 540
DB	481	IIDDSITIIKVKYQDGNQLPDSRIYDYSEYEDVTHDDYAQLGNNDVNIINFGNIDSPYI 540
QY	541	IKVISKYPNKKDDVTITQQVTWQTTINEYTGFRFASVNTTAFSTSSGQGGDLPPK 600
DB	541	IKVISKYPNKKDDVTITQQVTWQTTINEYTGFRFASVNTTAFSTSSGQGGDLPPK 600
QY	601	TYKIGDYVWEDVDKQIGQNTNDNEKPLSNVLVLTYPDGTGSKSVRTDEDKGYQFDGLKNG 660
DB	601	TYKIGDYVWEDVDKQIGQNTNDNEKPLSNVLVLTYPDGTGSKSVRTDEDKGYQFDGLKNG 660
QY	661	LTWKITFETPEGTPTLKHSGTTPALDSSEGNVWVITINGQDDMTIDSGFYQTPKYSIGNY 720
DB	661	LTWKITFETPEGTPTLKHSGTTPALDSSEGNVWVITINGQDDMTIDSGFYQTPKYSIGNY 720
QY	721	VWYDTNKGIGQDEKIGISGVKVTLDKNGNIIITTTTDENGKYQPDNLNGNVIYVHFDK 780
DB	721	VWYDTNKGIGQDEKIGISGVKVTLDKNGNIIITTTTDENGKYQPDNLNGNVIYVHFDK 780
QY	781	PSGMTQTITDSDGDEQADGEEHVHTITDHDDFSINDGYDDSDSDSDSDSDSDSDSDSD 840
DB	781	PSGMTQTITDSDGDEQADGEEHVHTITDHDDFSINDGYDDSDSDSDSDSDSDSDSDSD 840
QY	841	SD 900
DB	841	SD 900
QY	901	SD 960
DB	901	SD 960
QY	961	SD 1020
DB	961	SD 1020
QY	1021	SDSGSDSDSDSDSDSDSDSDSDSDSGNSDKSTKDKLPDTGANEDYGSKGTLLGTLFAGLGALL 1080
DB	1021	SDSGSDSDSDSDSDSDSDSDSDSDSGNSDKSTKDKLPDTGANEDYGSKGTLLGTLFAGLGALL 1080
QY	1081	LGKRRKRNKN 1092
DB	1081	LGKRRKRNKN 1092
RESULT 3		
ABP40469		
ID	ABP40469	standard; protein; 930 AA.
XX	AC	ABP40469;
XX	DT	24-JUL-2002 (first entry)
XX	DE	Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:5314.
XX	KW	Staphylococcus epidermidis; open reading frame; ORF; bacterial infection; antibacterial; gene therapy.
XX	OS	Staphylococcus epidermidis.
XX	PN	US6380370-B1.
XX	PD	30-APR-2002.
XX	PF	13-AUG-1998; 98US-00134001.
XX	PR	14-AUG-1997; 97US-0055779P.
XX	PR	08-NOV-1997; 97US-0064964P.
XX	PA	(GENO-) GENOME THERAPEUTICS CORP.
XX	PI	Doucette-Stamm LA, Bush D;
XX	DR	WPI; 2002-381255/41.
XX	DR	N-PSDB; ABN93014.
PT	Novel isolated nucleic acid encoding a Staphylococcus epidermis polypeptide, useful for diagnosing and treating bacterial infections.	
XX	PS	Disclosure; SEQ ID NO 5314; 267pp; English.
XX	CC	ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading frame (ORF) nucleic acid sequences which encode the amino acid sequences given in ABP35124 to ABP37960. The S. epidermidis sequences have antibacterial activity and can be used in gene therapy. The sequences can also be used in the diagnosis and treatment of bacterial infections, particularly S. epidermidis infections. The sequences can be used to screen for compounds able to interfere with the S. epidermidis life cycle or inhibit S. epidermidis infection. N.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the USPTO web site
XX	SQ	Sequence 930 AA;
Query Match		79.5%; Score 4490; DB 5; Length 930;
Best Local Similarity		81.2%; Pred. No. 2e-210;
Matches 884; Conservative 19; Mismatches 26; Indels 160; Gaps 3;		
QY	4	KKNLLTKKPIANKSNKYAIRFTVGTASIVIGATLLFGLGNEAKAENSVDVKDSN 63
DB	2	KKNLLTKKPIANKSNKYAIRFTVGTASIVIGATLLFGLGNEAKAENSVDVKDSN 61
QY	64	TDDLSNDSDQSSDEEKNDVINNNQSIINTDDNNQIIKKEETNNYDGIKRSERDTESTN 123
DB	62	MDELSDNSDQSSNEEKNDVINNNQSIINTDDNNQ- IKKEETNDALENRSKQITQSTTN 120
QY	124	VDENEATFLQKTPQDNTLHTEEEVKSSSVSSSSIDTAQPSHTTINREESVQTS 183
DB	121	VDENEATFLQKTPQDNTLHTEEEVKSSSVSSSSMDTAQPSHTTINREESVQTS 19



Db 214 ENRVSDFANSKILIESNTESKENTIEQNKVREDSITSQPSYKNIDEKISNQDELLN 273  
Qy 244 LPINEYENKARPLSTTSAPSIKRVTNQLAARQGSNVNHLIKVTDQSITEGYDDSEGI 303  
Db 274 LPINEYENKVRPLSTTSAPSSSRVTNQLAARQGSNVNHLIKVTDQSITEGYDDSGII 333  
Qy 304 KAHAENLIYDVTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFTIKIKDNGSEIIATGT 363  
Db 334 KAHAENLIYDVTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFAIKIKDNGSEIIATGT 393  
Qy 364 YDNKNKQITVTFDYDKYENIKAHKLKTSYIDKSKVPNNNTKLDVEYKTSALSSVNKTIT 423  
Db 394 YDNKNKQITVTFDYDKYENIKAHKLKTSYIDKSKVPNNNTKLDVEYKTSALSSVNKTIT 453  
Qy 424 VEYQRPENRNTANLQSMFTNIDTKNHTVEQIYINPLYSAKETNVIISGNGDEGSTIID 483  
Db 454 VEYQRPENRNTANLQSMFTNIDTKNHTVEQIYINPLYSAKETNVIISGNGDEGSTIID 513  
Qy 484 DSTIIVKYKVDGNQNLPSNRRIYDYSEYEDVTNDYDQAOLGNNDVNFNIDSPYIIKV 543  
Db 514 DSTIIVKYKVDGNQNLPSNRRIYDYSEYEDVTNDYDQAOLGNNDVNFNIDSPYIIKV 573  
Qy 544 ISKYDPNKDYYTIIQOTVTMOTTTINEYTGERTASVDNTIAPSTSSGQGGDLPPPKTYK 603  
Db 574 ISKYDPNKDYYTIIQOTVTMOTTTINEYTGERTASVDNTIAPSTSSGQGGDLPPPKTYK 633  
Qy 604 IGYVWEDVDKQIGQNTDNEKPLSNVLTLYPDGTSKSVRTDEGKYQFDGLKNGLTY 663  
Db 634 IGYVWEDVDKQIGQNTDNEKPLSNVLTLYPDGTSKSVRTDEGKYQFDGLKNGLTY 693  
Qy 664 KITFETPEGTPTLKHSQTNPALDSEGNVWVTINGQDDMTIDSGFYQTPKYSGLNYVWY 723  
Db 694 KITFETPEGTPTLKHSQTNPALDSEGNVWVTINGQDDMTIDSGFYQTPKYSGLNYVWY 753  
Qy 724 DTNKGQIGQDEKIGISGVKVTLDENGNIIISTTTTIDENGKYQFDNLNSGNYIVHFKPSPG 783  
Db 754 DTNKGQIGQDEKIGISGVKVTLDENGNIIISTTTTIDENGKYQFDNLNSGNYIVHFKPSPG 813  
Qy 784 MTQTTTDSGDDQDADGEEVHVITTDHDDFSIDNGYDDSDSDSDSDSDSDSDSDSDS 843  
Db 814 MTQTTTDSGDDQDADGEEVHVITTDHDDFSIDNGYDDSDSDSDSDSDSDSDSDSDS 866  
Qy 844 DS 903  
Db 867 ----- 866  
Qy 904 DS 963  
Db 867 ----- 866  
Qy 964 DS 1023  
Db 867 DS 913  
Qy 1024 GSDS 1083  
Db 914 -----NSSDKNTKDKLPDTCANEDHSDSGTLLGALPAGLILGK 954  
Qy 1084 RKNRKNKN 1092  
Db 955 RKNRKNKN 963

## RESULT 5

AA70120

ID AA70120 standard; protein; 991 AA.

XX AC

XX AC

XX AC

DT 06-JUN-2000 (first entry)

XX XX

DE Staph. epidermidis serine-aspartate repeat region protein SdrG.

XX Multicomponent vaccine; immunostimulatory; antibacterial; MSCRAMM;  
KW microbial surface components recognising adhesive matrix molecules;  
KW collagen binding protein; CBP; CNA; fibrinogen binding protein;  
KW Clumping factor A; ClfA; Clumping factor B; ClfB; FnBP;  
KW fibronectin binding protein; Staphylococcus infection;  
KW serine-aspartate repeat region protein; SDR protein; SdrG.  
XX Staphylococcus epidermidis.  
OS  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 14 /note= "Encoded by in-frame stop codon TAG"  
FT Misc-difference 33 /note= "Encoded by in-frame stop codon TGA"  
FT Misc-difference 964 /note= "Encoded by in-frame stop codon TAA"  
FT Misc-difference 980 /note= "Encoded by in-frame stop codon TAG"  
FT Misc-difference 989 /note= "Encoded by in-frame stop codon TAA"  
FT Misc-difference 991 /note= "Encoded by in-frame stop codon TAA"  
PN WO200012131-A1.  
XX  
XX 09-MAR-2000.  
XX  
PF 31-AUG-1999; 99WO-US019727.  
XX  
XX 31-AUG-1998; 98US-0098439P.  
XX  
XX (INHI-) INHIBITEX INC.  
XX (TEXA) UNIV TEXAS A & M SYSTEM.  
XX (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.  
XX  
XX Patti JM, Foster TJ, Hook M;  
XX WPI: 2000-237781/20.  
XX N-PSDB; AA251202.  
XX  
XX Composition used for generating immune response or for inhibiting  
PT microbial colonization in an animal comprises antibodies that bind  
PT collagen binding protein, fibrinogen binding protein and, optionally,  
PT fibronectin binding protein.  
XX  
XX Claim 8; Fig 4; 115pp; English.  
XX  
XX The patent discloses multicomponent vaccines containing selected  
CC combinations of bacterial binding proteins termed MSCRAMM (microbial  
CC surface components recognising adhesive matrix molecules) or their  
CC antibodies. A vaccine composition is provided that includes collagen  
CC binding protein or peptide, e.g. CNA, a fibrinogen binding protein  
CC preferably Clumping factor A (ClfA) or Clumping factor B (ClfB), and  
CC optionally a fibronectin binding protein e.g. FnBP-A. The vaccines are  
CC useful for imparting protection against a broad spectrum of  
CC staphylococcal strains and for inhibiting microbial colonisation.  
CC especially of Staphylococcus aureus, in an animal. The combinations can  
CC also be used to select donor blood pools for the preparation of purified  
CC blood products for passive immunisation. The present sequence is a series  
CC -aspartate repeat region protein, SdrG from Staphylococcus epidermidis.  
CC The Sdr protein is useful in vaccine preparation in combination with  
CC specific bacterial binding proteins. These vaccines can be used to treat  
CC a broad spectrum of bacterial infections, including those arising from  
CC both coagulase-positive and coagulase-negative bacteria  
XX  
XX Sequence 991 AA;

Query Match 79.4%; Score 4485; DB 3; Length 991;

Best Local Similarity 81.1%; Pred. No. 3.7e-210;

Matches 883; Conservative 19; Mismatches 27; Indels 160; Gaps 3;

Qy 4 KKNLLTKKKPIANKSNKYAIRFTVTGTASIVIGATLLFGLGHNEAKAESNVQDVKDSN 63

Db 35 KKNLLTKKKPIANKSNKYAIRFTVTGTASIVIGATLLFGLGHNEAKAESNVQDVKDSN 94





CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 670 AA;

Query Match 59.4%; Score 3351.5; DB 6; Length 670;  
Best Local Similarity 96.3%; Pred. No. 3.4e-155;  
Matches 648; Conservative 0; Mismatches 2; Indels 23; Gaps 2;

Qy 440 MFTNIDTKHVEQTIYINPLRYSAKETNWNISGNDEGSTIIDSTIIKVKYKVDNQNL 499  
Db 1 MFTNIDTKHVEQTIYINPLRYSAKETNWNISGNDEGSTIIDSTIIKVKYKVDNQNL 60

Qy 500 PDSNRIYDSEYEDVNDVAQLGNNDVNFNFGNIDSPYIIKVKYKVDNPKDDYTTIQ 559  
Db 61 PDSNRIYDSEYEDVNDVAQLGNNDVNFNFGNIDSPYIIKVKYKVDNPKDDYTTIQ 120

Qy 560 TVTMTTINEYTGFRFASVDNTIAPSTSSGQGGDLPPKTYKIGYVWEDVDKGIQN 619  
Db 121 TVTMTTINEYTFE---ASYDNTIAFSTSGQGGDLPPKTYKIGYVWEDVDKGIQN 177

Qy 620 TNDNEKPLSNVLTLYPDGTSKSVRTDEGKYQFDGLKGLTYKITFETPEGYTPTLKH 679  
Db 178 TNDNEKPLSNVLTLYPDGTSKSVRTDEGKYQFDGLKGLTYKITFETPEGYTPTLKH 237

Qy 680 SGTNPALDSGNSVWVTVINGQDDMTIDSGFYQTPKYSGLGNVWYDTNKGIGQDDEKGIS 739  
Db 238 SGTNPALDSGNSVWVTVINGQDDMTIDSGFYQTPKYSGLGNVWYDTNKGIGQDDEKGIS 297

Qy 740 GVKVTLKDENGNIISTTTTDENGKYQFDNLNSGNYIVHFDKPSGWTQTTTDSGDDDEQDA 799  
Db 298 GVKVTLKDENGNIISTTTTDENGKYQFDNLNSGNYIVHFDKPSGWTQTTTDSGDDDEQDA 357

Qy 800 DGEVHVITTDHDDFSDNGYVDE-----SDSDSDSDSDSDSDSDSDSDSDSDS 839  
Db 358 DGEVHVITTDHDDFSDNGYVDESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 417

Qy 840 DS 899  
Db 418 DS 477

Qy 900 DS 959  
Db 478 DS 537

Qy 960 DS 1019  
Db 538 DS 597

Qy 1020 DS 1079  
Db 598 DS 657

Qy 1080 LLGKRRKRNKVN 1092  
Db 658 LLGKRRKRNKVN 670

RESULT 8  
ABM79020  
ID ABM79020 standard; protein; 549 AA.  
XX  
AC ABM79020;  
XX  
XX  
DT 15-JAN-2004 (first entry)  
XX  
XX Staphylococcus epidermidis polypeptide.  
DE Infection; antibacterial; vaccine.  
XX  
KW  
XX  
OS Staphylococcus epidermidis.

XX WO2003076470-A1.  
PN  
XX  
PD 18-SEP-2003.  
XX  
XX 05-MAR-2003; 2003WO-US006415.  
XX  
XX 05-MAR-2002; 2002US-0361324P.  
XX  
XX (INH1-) INHIBITEX INC.  
PA (TEXA) UNIV TEXAS A & M SYSTEM.  
XX  
XX Patti JM, Hutchins JT, Hall A, Domanski P, Patel P, Hook M;  
PI Robbins J, Vernachio J, Bowden MG;  
PI  
XX WPI; 2003-722324/68.  
XX  
XX New antibody recognizing a Staphylococcus epidermidis protein comprising  
PT SdrG N1N2N3, SdrG N2N3 or SdrGR2 useful for preparing a composition for  
PT treating or preventing a coagulase-negative Staphylococcal infection.  
XX  
XX Claim 20; Page 37; 78pp; English.  
XX  
XX The present sequence comprises the protein sequence of a polypeptide of a  
CC coagulase-negative Staphylococcus epidermidis. A claimed monoclonal  
CC antibody recognises this protein and is used in a claimed method of  
CC treating or preventing a coagulase-negative staphylococcal infection in a  
CC human or animal, e.g. a nosocomial coagulase-negative staphylococcal  
CC infection in low birth weight infants  
XX  
SQ Sequence 549 AA;

Query Match 49.4%; Score 2787; DB 7; Length 549;  
Best Local Similarity 98.7%; Pred. No. 7.9e-128;  
Matches 542; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 52 EENSVDQVKDSNTDDELSDNSDQSDDEKNDVINNNQSIINTDNNQIIKKEETNNYDGE 111  
Db 1 EENSVDQVKDSNTDDELSDNSDQSDDEKNDVINNNQSIINTDNNQIIKKEETNNYDGE 60

Qy 112 KSESDRTSTTVNDENEATFLQKTPQDNTHLTTEEVKESSESSVESNSSIDTAAQPSHTTI 171  
Db 61 KSESDRTSTTVNDENEATFLQKTPQDNTHLTTEEVKESSESSVESNSSIDTAAQPSHTTI 120

Qy 172 NREESVQTSNDVEDSHVSDPFANSKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNI 231  
Db 121 NREESVQTSNDVEDSHVSDPFANSKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNI 180

Qy 232 DEKISNQDELMLPINEYENKARPLSTTSAQPSIKRVTNVNQLAAPQGSNVNHLIKVTQDS 291  
Db 181 DEKISNQDELMLPINEYENKARPLSTTSAQPSIKRVTNVNQLAAPQGSNVNHLIKVTQDS 240

Qy 292 ITEGYDDSEGVIKAHDAENLIYDVTFFVDDKVKSGDTMTVIDKNTVPSDLTDSFTIPKI 351  
Db 241 ITEGYDDSEGVIKAHDAENLIYDVTFFVDDKVKSGDTMTVIDKNTVPSDLTDSFTIPKI 300

Qy 352 KNSGEIATGYDNKNKQITVTFDYDVKYENIKAHKLTSYIDKSKVPNNNTKLDVEY 411  
Db 301 KNSGEIATGYDNKNKQITVTFDYDVKYENIKAHKLTSYIDKSKVPNNNTKLDVEY 360

Qy 412 KTALSSVNKTIIVVEYQRPENRNTANLQSMFTNIDTKHTEVEQTIYINPLRYSAKETNVI 471  
Db 361 KTALSSVNKTIIVVEYQRPENRNTANLQSMFTNIDTKHTEVEQTIYINPLRYSAKETNVI 420

Qy 472 SNGDGESTIIDSDSTIIKVKYKVDNQNLPSNRIYDSEYEDVNDVAQLGNNDVNFN 531  
Db 421 SNGDGESTIIDSDSTIIKVKYKVDNQNLPSNRIYDSEYEDVNDVAQLGNNDVNFN 480

Qy 532 FGNIDSPYIIKVKISYDNPDKDDYTTIQOTVTMTTINEYTGFRFASVDNTIAPSTSSGQ 591  
Db 481 FGNIDSPYIIKVKISYDNPDKDDYTTIQOTVTMTTINEYTGFRFASVDNTIAPSTSSGQ 540

Qy 592 QGGDLPPPEK 600



Db 541 GQGLPPEK 549

RESULT 9

AAV08643

ID AAV08643 standard; protein; 1166 AA.

AC AAV08643;

XX 20-MAR-2003 (revised)

DT 09-AUG-1999 (first entry)

XX XX

DE S. aureus SdrE protein.

XX Fibrinogen-binding protein; alpha chain; beta chain; ClfB; SdrC; SdrD; SdrE; fibrinogen; medical device; competitive inhibitor; pharmaceutical; treatment; infection; septicemia; osteomyelitis; mastitis; endocarditis; extracellular matrix; vascular graft; vascular stent; vaccine; intravenous catheter; artificial heart valve; cardiac assist device; antibacterial.

XX Staphylococcus aureus.

OS WO9927109-A2.

XX 03-JUN-1999.

XX 25-NOV-1998; 98WO-US025246.

XX 26-NOV-1997; 97US-0066815P.

PR 31-AUG-1998; 98US-0098427P.

XX (INH-) INHIBITEX INC.

PA (FORP-) FORP T/A BIORESEARCH IRELAND.

PA (TEXA) UNIV TEXAS A & M.

PA (PATTI) PATTI J M.

PA (JOSE) JOSEFFSON E.

PA (EIDH) EIDH D N.

PA (HOOK) HOOK M A O.

PA (PERK) PERKINS S E.

XX Patti JM, Foster TJ, Josefsson E, Eidhin DN, Hook MAO; Perkins SE;

PI WPI: 1999-357844/30.

XX N-PSDB; AAX77594.

DR Staphylococcus aureus fibrinogen-binding proteins for treating septicemia, osteomyelitis, mastitis or endocarditis.

PT Claim 8; Fig 9; 143pp; English.

XX This invention describes novel Staphylococcus aureus fibrinogen-binding proteins that bind both the alpha and beta fibrinogen chains. The proteins (and their encoding nucleic acids) are ClfB, SdrC, SdrD and SdrE. Staphylococcus aureus is thought to utilize fibrinogen to adhere to medical devices, binding proteins that bind both the alpha and beta fibrinogen chains (ClfB, SdrC, SdrD and SdrE) can therefore be used as competitive inhibitors to block this binding. Antibodies against ClfB, SdrC, SdrD and SdrE inhibit ClfB, SdrC, SdrD and SdrE mediated binding. The proteins of the invention can be used in a pharmaceutical composition for the treatment of Staphylococcus aureus infection e.g. septicemia, osteomyelitis, mastitis or endocarditis or to inhibit the binding of S. aureus to the extracellular matrix. The proteins or their fragments may be used to coat a medical device to reduce the S. aureus infection of an indwelling medical device, especially where the medical device is selected from the group consisting of vascular grafts, vascular stents, intravenous catheters, artificial heart valves, and cardiac assist devices. ClfB, SdrC, SdrD, SdrE, or an active fragment, subdomain or encoding gene may be used as a vaccine. The DS (aspartate serine) repeat region or a gene encoding it may be used as an identifying probe for the

CC identification of genes and encoding proteins from Staphylococcus aureus (other than CifA), S. hemolyticus, S. lugdenensis, and S. schleriferi.

CC The proteins of the invention have antibacterial activity. (Updated on 20 -MAR-2003 to correct PA field.)

XX

SQ Sequence 1166 AA;

Query Match 47.8%; Score 2698.5; DB 2; Length 1166;

Best Local Similarity 48.2%; Pred. No. 4.1e-123;

Matches 595; Conservative 158; Mismatches 267; Indels 215; Gaps 26;

QY 1 MINKKN-LLTKKKPIANKSNKYARKFTVGTASIVIGATLLPGLGHNKAKAENSVDV 59

DB 1 MINRDNKKAITKKGWISNRLNFKSIRKTYVTGASILVGLTFLPGLNQAKAAENT--ST 58

QY 60 KDSNTDELSDNDSSDEEKNDVINNQSINTDNNQIIEKEETNNYDGIERSDRT 119

DB 59 ENAKQDDATSDNKEVSVETENNSTNTP-----IKKE--TNTSQPEAKKETS 110

QY 120 STTNVDENEATFLQKTPQDNTHLTBEVEKES---SSVESNSISIDTAQPSHTTIREES 176

DB 111 SSTQKQNNVTATTTETKPN--IEKENVPSTDKTATEDSVILEKKAPNNT--NNDVT 166

QY 177 VQTSNVEDSHVSDPANSKIKESNT---ESGKEENTIEOPNKVKEDSTTSQPSGYTNIDE 233

DB 167 TKPS-----TSEPSTSEIQTKPTTPQBSTNIENSQPQPTPSKVD---NQVTDATNPKE 216

QY 234 KIS-NODELLNLP-----INEYENKARPLSTTSAQPSIKRV-----TYNQLAEEQ 277

DB 217 PVNVKEELKNPEKELVRLNDSNTDHTKPVATAPTSAVKRVNAKRFVAQAAPAAVA 276

QY 278 GSNVHLIKVTDQSIPEGVDSEGVKAHAENLIYDVTFEVDDKVKSGDTMTVDIDKNT 337

DB 277 SNNVNDLIKVTQTIKVG-DGKDNVAAAHGDKIEYDTEFTIDNKVKKGDTMTINDKVN 335

QY 338 VPSDLTDSFTIPKIDNSGEIATGTYDNKQIYTTFTDYDVKYENIAHLKLTSLVIDK 397

DB 336 IPSDLTDRKNDPIDITDPSGEVIAKGFDRATQIYTTFTDYDVKYEDIKSLRLTLYSIDK 395

QY 398 SKVPNNNTKLDVEYKTALSSVNKTITVEYQRPNNRNTANLQSMFTNIDTKNHTVEQTIYI 457

DB 396 KIVP-NETSLNLTFFATAGKETSQNVTVQDPVHGDGSIQSIFFTKLDEKQIEQIYV 454

QY 458 NPLRYSAKETNNVNSNG-----DEGSTIIDSTIIKVKYKGDNQNLPSNRIYD 508

DB 455 NPLKKSATNTKVDIAGSQVDYGNIKLGNSTIIDQNTBIKVKVNSDQOLPQSNRIYDF 514

QY 509 SEYEDVTND-DYAQLGNNDNVNFGNIDSPYIIKVISKYDKNKDYTTIQQVTTMOTTI 567

DB 515 SQYEDVTISQFNNKSFNNVATLDFGDIINSAYIIKVVSKYTPSTSGELDTAQGTSMTT- 573

QY 568 NEYTGFEFTASYDNTIAFSTSSQGGQGLD-PPEKTYKIGDYVWEDVDKQIQNTNNEKP 626

DB 574 DKY-GYNTAGVSNFVTSNDTGGGDTVKPEKLYKIGDYVWEDVDKGVQGTDSKEP 632

QY 627 LSNVLTLYTPDGTSKSVRTDEBGKQFPLGNKGLYKITYFTPEPEGYTFLPKHSGNPAL 686

DB 633 MANVLVLTYPDGTSTKSVRTDANGHYEFGGLDGETYTVKFTPTGYLPTKVGNTGDGEK 692

QY 687 DSEGNVWVTTINGQDDMTIDSGFYQTPKYSLGNVYVYDNTKQIGOGDDEKGISGVKVTLK 746

DB 693 DSGSSVTVKINGKODMSLDTGFYKPKYNLGDYVWEDNKGQIQDANEPGIKDVKVTLK 752

QY 747 DENGNIIISITT----- 757

DB 753 DSTGKVIQITTTDASGKYKFTDLNNGYTFETPAGYTPVKNVTTADDKDSNGLTTGV 812

QY 758 ----- 757

DB 813 IKDADNMVTLDRGFKTPKYSKSLGDYVWYVNSKQKQDSTEKGIKDVTVTLQNEKGEVIGTT 872

QY 758 -TDENKQYQFDNLNSGNYIVHFDKPSGMTQTTTDSGDDDEQDADGGEVHVHTTIDHDFS 816



Db 816 ADMWTLDRGYKTPKYSGLDYVWDSNKGQSTEGIKVDVTVTLQNEKGEVIGTKTD 875  
 Qy 760 ENGKYQFDNLNSGNYIVHFDKPSGWTQTTTDSGDDQDQADGEEVHVHTIDHDDFSDNG 819  
 Db 876 ENGKYRFDNLDSGKYVIFPKPAGLITQVNTTDDO-KDADGGEVDVTTIDHDDFTLNG 934  
 Qy 820 YDDESDS 879  
 Db 935 YFEEDT-----SDS 983  
 Qy 880 DS 939  
 Db 984 DS 1043  
 Qy 940 DS 999  
 Db 1044 DS 1103  
 Qy 1000 DS 1058  
 Db 1104 DSDAGKHTPVKPM-----TTKDHNNKAKALPETGS 1134  
 Qy 1059 NEDYGSKGTILGTLFAGLGALL-GKRRKNK 1089  
 Db 1135 ENNGSNATLFGGLFAALGSLLLFGRKKQNK 1166  
 RESULT 11  
 ID ABM79015 standard; protein; 560 AA.  
 AC ABM79015;  
 XX  
 XX 15-JAN-2004 (first entry)  
 DT Staphylococcus epidermidis SdrG N1N2N3 domain.  
 DE  
 XX SdrG; surface protein; infection; antibacterial; vaccine.  
 XX Staphylococcus epidermidis.  
 OS Staphylococcus epidermidis.  
 XX W02003076470-A1.  
 XX 18-SEP-2003.  
 XX 05-MAR-2003; 2003WO-US006415.  
 XX 05-MAR-2002; 2002US-0361324P.  
 XX (INH1-) INHIBITEX INC.  
 XX (TEXA) UNIV TEXAS A & M SYSTEM.  
 XX Patti JM, Hutchins JT, Hall A, Domanski P, Patel P, Hook M;  
 XX Robbins J, Vernachio J, Bowden MG;  
 XX WPI; 2003-722324/68.  
 XX N-PSDB; ACF80624.  
 XX New antibody recognizing a Staphylococcus epidermidis protein comprising  
 PT SdrG N1N2N3, SdrG N2N3 or SdrGTR2 useful for preparing a composition for  
 PT treating or preventing a coagulase-negative Staphylococcus infection.  
 XX  
 XX Claim 27; Page 24-25; 78pp; English.  
 XX The present sequence comprises the protein sequence of the N1N2N3 region  
 CC (amino acids 50-597), or putative A domain, of the SdrG surface protein  
 CC of coagulase-negative Staphylococcus epidermidis. A claimed antibody  
 CC recognises a protein selected from SdrG N1N2N3, N2N3 and TR2. The  
 CC antibody may be a monoclonal antibody, including a chimeric, murine,  
 CC humanized, human or single chain monoclonal antibody, which prevents a  
 CC coagulase-negative staphylococcus infection in a human or animal by  
 CC inhibiting binding of staphylococcus bacteria to fibrinogen. Such

CC antibodies can be used to treat or prevent staphylococcal infections  
 CC including nosocomial coagulase-negative staphylococcal infections in low  
 CC birth weight infants. A claimed vaccine comprises the isolated SdrG  
 CC N1N2N3, N2N3 or TR2 protein  
 XX  
 SQ Sequence 560 AA;  
 Query Match 46.5%; Score 2624.5; DB 7; Length 560;  
 Best Local Similarity 92.3%; Pred. No. 6.6e-120;  
 Matches 512; Conservative 17; Mismatches 25; Indels 1; Gaps 1;  
 Qy 46 HNEAKAEENSVDKSDNTDDELSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 105  
 Db 7 HHHGSEENVQVQKSDNMDDELSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 65  
 Qy 106 NYDGIKRESDERTSTNVNDENEATFLQKTPQDNTLHTEEVKESSESSSSSIPDAQQ 165  
 Db 66 SDAENIRSKDITQSTTNVDENEATFLQKTPQDNTLHTEEVKESSESSSSSMDTAQQ 125  
 Qy 166 PSHTTINRESVQTSNVEDSHVSDPANSKIKESNTESGKEENTIEQPNKVKEDSTTSQP 225  
 Db 126 PSHTTINSEASIQTSNEENSRVSDPANSKIKESNTESGKEENTIEQPNKVKEDSTTSQP 185  
 Qy 226 SGYTNIDEKISNQDELLNLPINEYENKARPLSTTSQAQPSIKRVTVNQLAAEQGSNNVHLI 285  
 Db 186 SSYKNIDEKISNQDELLNLPINEYENKARPLSTTSQAQPSIKRVTVNQLAAEQGSNNVHLI 245  
 Qy 286 KVTQDSITEGYDDSEGVKAKHAENLIYDVTPEVDDKVKSGDTMTVDIDKNTVPSDLTDS 345  
 Db 246 KVTQDSITEGYDDSGIKAHAENLIYDVTPEVDDKVKSGDTMTVDIDKNTVPSDLTDS 305  
 Qy 346 FPIPKIKONGELIATGTYDNKQKQITTYTFTDYDKYENIKAHKLKLTSTYIDKSKVPPNNT 405  
 Db 306 FAIPKIKDNGSEIATGTYDNKQKQITTYTFTDYDKYENIKAHKLKLTSTYIDKSKVPPNNT 365  
 Qy 406 KLDVEYKTALSSVNKTIITVEYQRPENRTANLQSMFTNIDTKHTEQTIYINPLYSK 465  
 Db 366 KLDVEYKTALSSVNKTIITVEYQRPENRTANLQSMFTNIDTKHTEQTIYINPLYSK 425  
 Qy 466 ETNNISGNGDEGTTIIDSTIIKVKYKVDGNQNLPSNRIYDYSEYEDVTNDDYAQLGNN 525  
 Db 426 ETNNISGNGDEGTTIIDSTIIKVKYKVDGNQNLPSNRIYDYSEYEDVTNDDYAQLGNN 485  
 Qy 526 NDVNFNGNIDSPYIIKVKSKYDPNKKDYYTTOQTVMQTTINSEYTGEPRTASYDNTIAF 585  
 Db 486 NDVNFNGNIDSPYIIKVKSKYDPNKKDYYTTOQTVMQTTINSEYTGEPRTASYDNTIAF 545  
 Qy 586 STSSGGQGGDLPEK 600  
 Db 546 STSSGGQGGDLPEK 560  
 RESULT 12  
 ID ABU42327  
 ID ABU42327 standard; protein; 1141 AA.  
 AC ABU42327;  
 XX  
 XX 19-JUN-2003 (first entry)  
 DT  
 XX  
 XX Protein encoded by Prokaryotic essential gene #27854.  
 DE Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 XX  
 XX Staphylococcus aureus.  
 OS  
 XX  
 XX W0200277183-A2.  
 XX 03-OCT-2002.  
 XX  
 XX 21-MAR-2002; 2002WO-US0009107.  
 XX  
 XX 21-MAR-2001; 2001US-00815242.  
 PR









antibodies. A vaccine composition is provided that includes collagen binding protein or peptide, e.g. CNA, a fibrinogen binding protein preferably Clumping factor A (ClfA), or Clumping factor B (ClfB), and optionally a fibronectin binding protein e.g. FnBP-A. The vaccines are useful for imparting protection against a broad spectrum of *Staphylococcal* strains and for inhibiting microbial colonisation, especially of *Staphylococcus aureus*, in an animal. The combinations can also be used to select donor blood pools for the preparation of purified blood products for passive immunisation. The present sequence is a serine-aspartate repeat region protein, SdrF from *Staphylococcus epidermidis*. The Sdr protein is useful in vaccine preparation in combination with specific bacterial binding proteins. These vaccines can be used to treat a broad spectrum of bacterial infections, including those arising from both coagulase-positive and coagulase-negative bacteria

Query Match	40.8%;	Score 2303.5;	DB 3;	Length 1802;
Best Local Similarity	34.6%;	Pred. No. 1.2e-103;		
Matches	605;	Conservative 146;	Mismatches 313;	Indels 685; Gaps 46;
Qy	2	INKKNLLTKKPKTANKSNKVAIRKFTVGTASIVIGATLLFGLGHNEAKAEN-----S	55	
Db	46	INKRVDF-----SNKNKYISIRKFTVGTASILVGATLMFGAADNEAKAAEDNQLFSAS	99	
Qy	56	VQDVKDS--NTDDBLS--DSNDQSDSDEKNDVINN-----NOSIN	91	
Db	100	KEEQKGRSDNENSKLVQDLONGSHSSEKTTNVNNAIEVKKVEAPTITDYSKPKANEAV	159	
Qy	92	TDD-----NNQIL--KKEETNNVDGIEK-----RSDRTSTNNVDEN	127	
Db	160	TNESTKPTTEAPTWNESIAETPKTTTQODSTTEKNPNSLKONLNSSTTSKESKTDH	219	
Qy	128	EATFLQ--KTQDNTHLTEEEVK--ESSSVESGNSIDTAQOPS-----	167	
Db	220	STKQAQMSTWKSNLTDWSTPQSEKTSQAANDSTDNQSAPSKQDLSKPSQKVYTKFN	279	
Qy	168	-----HTTIN--RESSVOTSDNVEDSHVSDFANSIKESNTBSGKEENTIEQPNKYKE	218	
Db	280	DEPTQDVEHTTTKLKTPSVSTDSVNDK--QDYTRSAV-----ASLGVDSNETEAITNAVR	333	
Qy	219	DSITTSQSPSGYNTDEKI-----SNODELLNLP--INEYENKARPLSTTSAQPSIKR	267	
Db	334	DNLDLKAASREQINEAIIAEBALKXDFNPDPYGVDTPLALNRSQSKNSP--HKSASP--R	388	
Qy	268	VTYNQLAAE--QGSNVNHLIKVTDQ--SITEGYDDSEGVIKAHDAENLIYDVTPVEDDKV	323	
Db	389	MNLSLAAEENPNSGKVVNDKVKIINPTLSLNKSNHANNVIWPTSNQFNLKANYELDDSI	448	
Qy	324	KSGDTMTVDIDKNVTPSDLTDSFTPIKIKONGSBI IATGVYDNKNKOITVYTFDQVYDKE	383	
Db	449	KEGDTFTIKYGOVIRPGGLEPAIKTQLRSKQSGSIVANGVYDKTTNTTFTVTVYDQYQ	508	
Qy	384	NIKAHLKLTSYIDKS KVPNNNTKLDVYKTKALS VNKTIIVEYQRPENETANLOSFTN	443	
Db	509	NITGSPDLIATPKRETAKONQVPMEVTIANEVVKKDFIDVYGNKKKDNNTT---AAVAN	565	
Qy	444	IDTKNHTVEQTIYINPLR-----YSAKETN-----	468	
Db	566	VDNVNNKHNEWYLVNQNQNPKYAKYFTVKNGEFPFGEVKVYEVTDTNAMVDSFNPLN	625	
Qy	469	-----VNISNGND-----EG	478	
Db	626	SSNVKDVTSQFAPKVSADGTRVDINFARSWANGKKYIVTQAVRPTGTGNVYTEYWLTRDG	685	
Qy	479	STIIDD-----STII-----KVYKVGD-----NON-----	498	
Db	686	TTWNDFYRGCTKSTVTYYLNGSSTAQGDNFTYSLGDVYWLDKXKXGVQDDDEKGLAGVYV	745	
Qy	499	LPDSN--RI-----YDYSEYEBDV-----NDDYAQGLGNVNDVN-----	529	
Db	746	TLKDSNNRELQVTVTDQSGHYQFDNLNGYTVTFVEFALPDNFTYSPANNSTDAIDSDGER	805	

Search completed: October 5, 2004, 20:48:21  
Job time : 138.638 secs

[illegible]



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:41:22 ; Search time 122.638 Seconds  
(without alignments)  
2809.455 Million cell updates/sec

Title: US-10-806-288-15  
Perfect score: 5646  
Sequence: 1 MINKNNLTGKKPIANKSN.....FAGLGALLGKKRKKNKN 1092

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : SPTREMBL.25.\*
- 1: sp\_archaea.\*
  - 2: sp\_bacteria.\*
  - 3: sp\_fungi.\*
  - 4: sp\_human.\*
  - 5: sp\_invertebrate.\*
  - 6: sp\_mammal.\*
  - 7: sp\_mhc.\*
  - 8: sp\_organelle.\*
  - 9: sp\_phase.\*
  - 10: sp\_plant.\*
  - 11: sp\_rodent.\*
  - 12: sp\_virus.\*
  - 13: sp\_vertebrate.\*
  - 14: sp\_unclassified.\*
  - 15: sp\_rvirus.\*
  - 16: sp\_bacteriaph.\*
  - 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5646	100.0	1092	2 070022	070022 staphylococ
2	5440	96.4	1056	16 08CQ72	08CQ72 staphylococ
3	4485	79.4	931	2 09K113	09K113 staphylococ
4	2756	48.8	1171	2 09KWX6	09KWX6 staphylococ
5	2698.5	47.8	1166	2 086489	086489 staphylococ
6	2589	45.9	1141	16 099W46	099W46 staphylococ
7	2589	45.9	1141	16 08NXX5	08NXX5 staphylococ
8	2586	45.8	1141	16 0932F7	0932F7 staphylococ
9	2344.5	41.5	1633	16 08CMP4	08CMP4 staphylococ
10	2303.5	40.8	1733	2 09K114	09K114 staphylococ
11	2234	39.6	1385	16 099W47	099W47 staphylococ
12	2093.5	37.1	953	16 099W48	099W48 staphylococ
13	2075.5	36.8	955	16 08NXX7	08NXX7 staphylococ
14	2037	36.1	1347	16 08NXX6	08NXX6 staphylococ
15	2023.5	35.8	1947	2 086487	086487 staphylococ
16	2006	35.5	1893	2 08KWM1	08KWM1 staphylococ

17	1880	33.3	1315	2 086488	086488 staphylococ
18	1652	29.3	989	16 099VJ4	099VJ4 staphylococ
19	1646.5	29.2	946	16 08NXX1	08NXX1 staphylococ
20	1628	28.8	933	2 053653	053653 staphylococ
21	1610	28.5	913	2 086476	086476 staphylococ
22	1581	28.0	935	16 0932C5	0932C5 staphylococ
23	1573	27.9	907	16 08NUL0	08NUL0 staphylococ
24	1536	27.2	877	16 099R07	099R07 staphylococ
25	1471	26.1	881	2 093MH7	093MH7 staphylococ
26	1216	21.5	1698	2 091C00	091C00 staphylococ
27	1203.5	21.3	3360	16 088XB6	088XB6 lactobacill
28	785	13.9	882	16 092AK9	092AK9 listeria in
29	782.5	13.9	903	16 08Y697	08Y697 listeria mo
30	747.5	13.2	2020	5 081D80	081D80 plasmodium
31	671.5	11.9	487	2 09K112	09K112 staphylococ
32	670.5	11.9	970	11 08VBY1	08VBY1 rattus norv
33	654.5	11.6	540	2 08KR21	08KR21 staphylococ
34	645.5	11.4	2310	16 08CMU7	08CMU7 staphylococ
35	640.5	11.3	485	16 08CNM7	08CNM7 staphylococ
36	639	11.3	579	13 07T138	07T138 brachydanio
37	624.5	11.1	2271	16 099QY4	099QY4 staphylococ
38	622	11.0	968	5 0815W9	0815W9 plasmodium
39	616.5	10.9	2275	16 08NNU3	08NNU3 staphylococ
40	612	10.8	2283	2 08VQ99	08VQ99 staphylococ
41	607	10.8	2900	5 0812E0	0812E0 plasmodium
42	565	10.0	1182	5 081D30	081D30 plasmodium
43	560	9.9	970	16 08DYL7	08DYL7 streptococc
44	554	9.8	3394	5 077384	077384 plasmodium
45	549	9.7	2570	2 092FF9	092FF9 streptococc

ALIGNMENTS

RESULT 1

070022 PRELIMINARY; PRT; 1092 AA.

AC 070022; 01-AUG-1998 (TREMREL. 07, Created)

DT 01-AUG-1998 (TREMREL. 07, Last sequence update)

DT 01-OCT-2003 (TREMREL. 25, Last annotation update)

DE Fibrinogen-binding protein precursor.

OS Staphylococcus epidermidis.

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

OX NCBI\_TaxID=1282;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=HB;

RX MEDLINE=98261511; PubMed=9596732;

RA Nilsson M., Frykberg L., Flock J.I., Pei L., Lindberg M., Guss B.;

RT "A Fibrinogen-binding protein of Staphylococcus epidermidis.";

RL Infect. Immun. 66:2666-2673(1998).

CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY AN AMIDE BOND (BY SIMILARITY).

DR EMBL, Y17116; CAA76638.1; -.

DR FIR, T30214; T30214.

DR GO; GO:0005618; C:cell wall; IEA.

DR GO; GO:0016020; C:membrane; IEA.

DR InterPro; IPR008454; Cna B.

DR InterPro; IPR005877; Gpos\_YSRK.

DR InterPro; IPR001899; Gram\_pos\_anchor.

DR Pfam; PF05738; Cna B; 2.

DR Pfam; PF00746; Gram\_pos\_anchor; 1.

DR Pfam; PF04650; YSRK signal; 1.

DR TIGRfams; TIGR01167; LPXTG anchor; 1.

DR TIGRfams; TIGR01168; YSRK signal; 1.

DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.

DR Cell wall; Peptidoglycan-anchor; Signal.

KW SIGNAL

FT CHAIN 1 51 POTENTIAL.

FT SEQUENCE 1092 AA; 119292 MW; 6542BC39AAD8B984 CRC64;

SQ

Query Match 100.0%; Score 5646; DB 2; Length 1092;



```
Db 421 TITVEQRPENRTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNWNISGDEGST 480
Qy 481 IIDSTIIKVKYKVDNQNLPSNRIDYSEYEDVTNDYAQLGNNDVNFNIGNIDSPYI 540
Db 481 IIDSTIIKVKYKVDNQNLPSNRIDYSEYEDVTNDYAQLGNNDVNFNIGNIDSPYI 540
Qy 541 IKVISKYPNKKDYTTIQQVTVMQTTINEYTGFEFTASVDNTIAFSTSGQGGLPPEK 600
Db 541 IKVISKYPNKKDYTTIQQVTVMQTTINEYTGFEFTASVDNTIAFSTSGQGGLPPEK 600
Qy 601 TYKIGDVWEDVDKGIQNTNDNEKPLSNVLVTLTPDGTSKSVRTDEDKYQFDGLKNG 660
Db 601 TYKIGDVWEDVDKGIQNTNDNEKPLSNVLVTLTPDGTSKSVRTDEDKYQFDGLKNG 660
Qy 661 LTYKITFETPEGYTPTLKHSNTNPDLDSEGSVWVTINGQDDMTIDSGFYQTPKYSLGNY 720
Db 661 LTYKITFETPEGYTPTLKHSNTNPDLDSEGSVWVTINGQDDMTIDSGFYQTPKYSLGNY 720
Qy 721 VYDTNKGDIQGDDEKIGSVKVTLDKENGNIISTTTTDENGKYQFDNLNSGNYIIVHFDK 780
Db 721 VYDTNKGDIQGDDEKIGSVKVTLDKENGNIISTTTTDENGKYQFDNLNSGNYIIVHFDK 780
Qy 781 PSQWTOITDSDGDDQDADGEEVHTIITDHDDFSIDNGYDDESDDSDSDSDSDSD 840
Db 781 PSQWTOITDSDGDDQDADGEEVHTIITDHDDFSIDNGYDDESDDSDSDSDSDSDSD 840
Qy 841 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 900
Db 841 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 900
Qy 901 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 960
Db 901 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 960
Qy 961 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1020
Db 961 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 988
Qy 1021 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1080
Db 989 ----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1044
Qy 1081 LGRKRNKRNKN 1092
Db 1045 LGRKRNKRNKN 1056

RESULT 3
Q9KI13 PRELIMINARY; PRT; 931 AA.
AC Q9KI13;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fibrinogen-binding protein SdrG.
GN SDRG.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K28;
RX MEDLINE=20340957; PubMed=10878118;
RA McCrea K W., Hartford O., Davis S., Ni Eidhin D., Lina G.,
RA Speziale P., Foster T.J., Hook M.,
RT "The serine-aspartate repeat (Sdr) protein family in Staphylococcus
epidermidis."
RL Microbiology 146:1535-1546 (2000).
CC -/- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY
ENBL; AF245042; AAF72510.1; -.
DR GO; GO:0005618; C:cell wall; IEA.
```

```
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR008454; Cna_B.
DR InterPro; IPR005877; Gpos_YSIK.
DR Pfam; PF05738; Cna_B; 2.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; YSIK_signal; 1.
DR TIGRfam; TIGR01167; LPXG_anchor; 1.
DR TIGRfam; TIGR01168; YSIK_signal; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 931 AA; 102955 MW; 591E657D97027116 CRC64;

Query Match 79.4%; Score 4485; DB 2; Length 931;
Best Local Similarity 81.1%; Pred. No. 6.4e-146;
Matches 883; Conservative 19; Mismatches 27; Indels 160; Gaps 3;

Qy 4 KKNLLTKKKPANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDQVKDSN 63
Db 3 KKNLLTKKKPANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENTVDQVKDSN 62
Qy 64 TDELSDSNDQSDDEKNDVINNNQSIINTDDNNQIIKKEETNNYDGIKRSERDSTTN 123
Db 63 MDELSDSNDQSSNEEKNDVINNSQSINTDDNQ-IKKEETNSDAIENRSDITQSTTN 121
Qy 124 VDENEATFLQKTPQDNTHLTEBEVKSSSVSSNSIDTAQOPSHHTTINREESVQTSNV 183
Db 122 VDENEATFLQKTPQDNTQLKEEVKSPSSVSSNSMDTAQOPSHHTTINSEASIQTSNE 181
Qy 184 EDSHVSDFPANSKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQDELLN 243
Db 182 ENSRVDFPANSKIIESNTESNKEENTIEQPNKVEDSIISQPSGYTNIDKISNQDELLN 241
Qy 244 LPINEYENKARPLSTTSAQPSIKRVTVNQLAAEQGSNNVHLIKVTPQSITEGYDDSEGI 303
Db 242 LPINEYENKVRPLSTTSAQPSKRVTVNQLAAEQGSNNVHLIKVTPQSITEGYDDSGII 301
Qy 304 KAHDAENLIYDTFVDDKVKSGDNTVDIDKNTVPSDLTDSFTIPIKIDNSGEIATGT 363
Db 302 KAHDAENLIYDTFVDDKVKSGDNTVNDKNTVPSDLTDSFPAIKIDNSGEIATGT 361
Qy 364 YNKNKQITVTFDYVDKYENIKAHKLTSYIDKSKVPNNNTKLDVEYKTLASVKNKTI 423
Db 362 YDNTNKQITVTFDYVDKYENIKAHKLTSYIDKSKVPNNNTKLDVEYKTLASVKNKTI 421
Qy 424 VEYQENENRTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNWNISGNGEGSTIID 483
Db 422 VEYQENENRTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNWNISGNGEGSTIID 481
Qy 484 DSTIIKVKYKVDNQNLPSNRIDYSEYEDVTNDYAQLGNNDVNFNIGNIDSPYIIV 543
Db 482 DSTIIKVKYKVDNQNLPSNRIDYSEYEDVTNDYAQLGNNDVNFNIGNIDSPYIIV 541
Qy 544 ISKYDPNKKDYTTIQQVTVMQTTINEYTGFEFTASVDNTIAFSTSGQGGLPPEKTYK 603
Db 542 ISKYDPNKKDYTTIQQVTVMQTTINEYTGFEFTASVDNTIAFSTSGQGGLPPEKTYK 601
Qy 604 IGDYVWEDVDKGIQNTNDNEKPLSNVLVTLTPDGTSKSVRTDEDKYQFDGLKNGLT 663
Db 602 IGDYVWEDVDKGIQNTNDNEKPLSNVLVTLTPDGTSKSVRTDEDKYQFDGLKNGLT 661
Qy 664 KITFETPEGYTPTLKHSNTNPDLDSEGSVWVTINGQDDMTIDSGFYQTPKYSLGNYV 723
Db 662 KITFETPEGYTPTLKHSNTNPDLDSEGSVWVTINGQDDMTIDSGFYQTPKYSLGNYV 721
Qy 724 DTNKGDIQGDDEKIGSVKVTLDKENGNIISTTTTDENGKYQFDNLNSGNYIIVHFDKPSG 783
Db 722 DTNKGDIQGDDEKIGSVKVTLDKENGNIISTTTTDENGKYQFDNLNSGNYIIVHFDKPSG 781
Qy 784 MTQTTTSDGDDQDADGEEVHTIITDHDDFSIDNGYDDESDDSDSDSDSDSDSDSDSDS 843
Db 782 MTQTTTSDGDDQDADGEEVHTIITDHDDFSIDNGYDDESDDSDSDSDSDSDSDSDSDS 834
```





RA	Cui L., Oguchi A., Aoki K.I., Nagai Y., Lian J., Ito T., Kanamori M.,	
RA	Matsumaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Ui Y.,	
RA	Takahashi N.K., Sawano T., Inoue R.I., Kaico C. S., Sekimizu K.,	
RA	Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,	
RA	Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,	
RA	Ogasawara N., Hayashi H., Hiramatsu K.,	
RT	"whole genome sequencing of methicillin-resistant <i>Staphylococcus</i>	
RT	<i>aureus</i> ";	
RL	Lancet 357:1225-1240 (2001).	
DR	EMBL; AP003131; BAB41752.1; -.	
DR	PIR; E89824; E89824.	
DR	GO; GO:0016020; C:membrane; IEA.	
DR	InterPro; IPR008454; Cna B.	
DR	InterPro; IPR005877; Gpos_Y5IRK.	
DR	InterPro; IPR001899; Gram_pos_anchor.	
DR	Pfam; PF05738; Cna B; 3.	
DR	Pfam; PF00746; Gram_pos_anchor; 1.	
DR	Pfam; PF04650; Y5IRK_signal; 1.	
DR	TIGRFAMs; TIGR01167; LPXTG_anchor; 1.	
DR	TIGRFAMs; TIGR01168; Y5IRK_signal; 1.	
DR	PROSITE; PS50847; GRAM_POS_ANCHORING; 1.	
KW	Complete proteome.	
QY	SEQUENCE 1141 AA; 124026 MW; 445419D08B8C5A4F8 CRC64;	

Query Match 45.9%; Score 2589; DB 16; Length 1141;  
 Best Local Similarity 46.7%; Pred. No. 4.3e-81;  
 Matches 577; Conservative 152; Mismatches 266; Indels 240; Gaps 27;

QY	1	MINKGN-NLLPKYPIANKSNKYAIRFTVGTASIVIGATLFLGLGHNPAKAENSVDV 59
DB	1	MINRDNKKAITKGMISNRLNKPFSIRKYTVGTASIIIVGLTFLFGLGNQAKAAENT--ST 58
QY	60	KDSNTDDDELSDNSQSDSEKNDVNNNQISINTDDNQIIKKEETNNYDGIKRSDEKTE 119
DB	59	ENAKQDDATTSDNKEVSEVENSTTENDTNP-----IKKE--TNTDSQPEAKEESTT 110
QY	120	STTVNDENEATFLQKTPQDNTHLTREEVKES---SSVESNSSIDTAQPSHTTINREES 176
DB	111	SSTQQQNNVTATTETKPN--TEKENVKPSTDKTATEDTSVLEBKAPNYT--NNDVT 166
QY	177	VQTSNVEDSHVDFANSKIKESNT---ESGKEENTIEOPNVKVEDSTTSQPSGYTNIDE 233
DB	167	TKPS-----TSEIQTKPTTPQESTNTIENSQPQTPPSKVD--NQVTDATNPKE 211
QY	234	KIS-NQDELLNLP-----INYEYENKARPLST--TSAQP-----SIKRVTVNQLAAEQ 277
DB	212	PVNVSKSELKNPPELKVRLVNDNNTDRSTKPVATPTSVAPRLNAKGRFAVAQPAVA 271
QY	278	GSNVNHLIKVTDQSITGEYDDSEGVKAHDAENLIYDVTFFVDDKVKSGDITMTVDIKNT 337
DB	272	SNVNVNDLITVKTIKVG-DGKNVAAAHGDKDIEYDTEFTIDNKKVGGDTWNTYDKNV 330
QY	338	VPSLDTSFTIPKIKDNGSIIATGIYDNKNKQITVTFDYDYDKYENIKAHKLKTSYIDK 397
DB	331	IPSLDTKNDPIDITDPSGSVIAGTFDKATQKITVTFDYDYDKYEDIKARLTLYSYIDK 390
QY	398	SKVPNNNTKLDVEYKTLALSSVWNTKITVEYORPNENRNTANLQSMFTNIDTKNHTVEQTIYI 457
DB	391	QAVP-NETSLNLTATAGKETSNQSVVDYQDPVWHDGNSIQSIIFTKLDENKQTIEQQIIV 449
QY	458	NPLRYSAKETNNVINSNG-----DEGSTIIDSTIIKYYKVGDNQNLPSDNRIYDY 508
DB	450	NPLKKTATNTKVDIAGSQVDYDNGIKLNGSTIIDQNTIEIKVYKVPNQQLPQSNRIYDF 509
QY	509	SEYEDVTND--DYAQLGNNDVNINFGNIDSPYIIKVISKIDPNKDDYTTIOQVTMTOTTI 567
DB	510	SYQEDVT-SQFDNKKSGNNVATLDFGDIINSAYIIKVISKYTPTSDBGELDIAQGTSMRTT- 568
QY	568	NEYTGEFTASYDNNTIAFSTSSGOGGLD--PREKTKIGDYVWEVDVKDGLONTNDNEKP 626
DB	569	DKY-GYNYAGYSNFIIVTNDTGGDGTVKPBEKLYKIGDYVWEVDVKDGVQGDSDSKEP 627
QY	627	LSNVLVTLVTPDGTSKSVRTDEBCKYQFDGLKNGLTIKTIFETPBGYVFTLKHGSGTNPAL 686

[illegible]



Dd	111	STQOQNNTAITETKPN--IEKENVKPSTDKTATEDTSVILEKKAPNT--NNDVT	167
Qy	177	VQTSNVEDSHVDSFANSIKESNT---ESGKEENTIEQNKVKEDSTTSQPSGYTNIDE	233
Dd	167	TKPS-----TSEIQTKPITPQESTWIENSQOPTSKVD---NQVTDATNPKE	211
Qy	234	KLS-NQDELLNLP-----INEYENKARPLST--TSAQP-----SIKRVTVNQLAAEQ	277
Dd	212	PVNVSKBELKNPEKLKELVRNDDNTDRSTKEVATAPTSAVKRLNAKRFVAPQAAVA	271
Qy	278	GSNVNHLIKVTDQSITEGVDDSEGVKHAHDAENLIYDVTFEVDVKSGDVTMTVIDKNT	337
Dd	272	SNVNDLITVTKQIKVG-DGKNVAADGKDIEYDTEFTIDNKKGDTMIINDKRV	330
Qy	338	VPSDLTDSFTPIKIKDNSEIIATGYDNKNKQIITYTFTDYVDKYENIKAHKLKLTYSIDK	397
Dd	331	IPSDITDRNDPDIITDPGSEVIAGTFDKATQKITVTFTDYVDKYEDIKARLTLYSIDK	390
Qy	398	SKVPNNNTKLDVEYKTAJLSSVNNKTLTVFVORENENRTANLOSMFNTIDTKHTVEQTIYI	457
Dd	391	QAVP-NEISLNLTFATAGKETSONVSVDYQDDPMVHGDSNIOISIFKLDENKQTIPEQQIYV	449
Qy	458	NPLRYSAKETNVNISGNG-----DEGSTIIDDSITIIKVKYKGDQNLPPDSNRIYDY	508
Dd	450	NPLKKTATNTKVDIAGSQVDYDGNIKLNGSTIIDQNTIEIKVKYKVPNQLPQSNRIYDF	509
Qy	509	SEYEDVTND-DYAOGLNNDNVINPGNTDSPIYIKVISKYDPNKDDYTTIOQVTVMQTTI	567
Dd	510	SOYEDVTSQFNKKSFSNNVATLDFGINDSAYIIKVKSYTPTSGELDIAQGTSMR'IT-	568
Qy	568	NEYTGEEFTASVDNTIAFSTSQGOQGL-PPEKTYKIGDYVWEDVDXGQONTNDNEKP	626
Dd	569	DKY-GYNYNAGYSNFI'VNTSDTGGDGVKPEEKLYKIGDYVWEDVDXGQGTDSKEKP	627
Qy	627	LSNVLVTLTYPDGTSKSVRTBEDGKYQPDGLKNGLTYKITTETPEGYPTPLKHSGTNPAL	686
Dd	628	MANVLVTLTYPDGTTKSVRTDANGHYEFGGLKDGETYTKFETPAGLYPTKVGNGTTDGEK	687
Qy	687	DSEGNVWVTINGOOD-----	702
Dd	688	DSNGSIITVKINGRDNDSLDTFYKPEKYNLGDYVWEDTNKDGIQDANEPGIKDVKVTLK	747
Qy	703	-----	702
Dd	748	DSTGKVGITTTTDSAGKYKFTDLNGYTVBFETPAGYPTPVKNTTAEDKDXDNGSLTTTGV	807
Qy	703	-----WTIDSGFYQTPKYSIGNVYVYDNTKDGIOGDDEKXIGSVKYVTLKXENGLI'ITT	756
Dd	808	IKADANWTLDSGFYKTPKYSIGDYVWDSNKGQDQDSTERGIKDVKVTLNKEGEVIGTT	867
Qy	757	TTDENGKYQFNLSNGNYIHVFDPKSGMTO'TTTDSGDDDEQADAGEEYHVHITDHDHFSI	816
Dd	868	KTDENGKRYRFDNLDSGKYKVIFEXKAPAGLTQVTVNTTEDD-KDADGGEVDVITDHDHDFIL	926
Qy	817	DNGYVYDDESDDSD	876
Dd	927	DNGYFESDT-----SD	955
Qy	877	SD	936
Dd	956	SD	1015
Qy	937	SD	996
Dd	1016	SD	1075
Qy	997	SD	1055
Dd	1076	SDSDSDAGKHTPVKPMs-----TTKOHNNKAKALPE	1106
Qy	1056	TGANEDYGSKGTLLGTLFAGLGALLL-GKRRKNRK	1089
Dd	1107	TGSENNNGSNATLFGGLFAGLGALLLFGRRKKQNK	1141







```
Qy 120 STTNVDEN-EATFLOKTPQDNTHLTBEEVKESSESSVSSSSIDTAQPSHTTINRESVQ 178
Db 103 TTSNGNKSIEKESVQSTGNGKVESTAKSDEQASPKSTNEDLTKQ-----TISNOEGLQ 157
Qy 179 TSDNVEDSHVSPFANSKIKESNTESGKEENTIQPN-KVKEDSTTQPSGYTNDISKIN 237
Db 158 -PDLLENKSVVN-----VQPTNEENKVKDAKTESTILNVKSDAICS-----NAETLVDN 205
Qy 238 QDELLNLPINEYENKARPLSTTSAQPSIKRVTVNQLAAEQGS-----NVNHLIKVTDQSI 293
Db 206 NSNSNN-----ENNADIILPKSTAPKSLNTRMAAIOQNSTSKNVNDLITSNTTLTV 259
Qy 294 EGYDDSEGLVKADAENLIYDVTFEVDDVKVSGDTWTDIDKTV-----PSDLTDSFT 347
Db 260 VDADNSKTIIPAQDYLKLSQIT--VDDKVKSGDYFTIKY-SDTQVQYGLNPEDIKN--- 313
Qy 348 IPKIKD-NSGEIATGYDNKNKQITVTFDYDYDKYENIKAKHLKLSYIDKSKVPNNNTK 406
Db 314 IGDIKDPNGETIATAKHTANLITYTFDYDRFNSVQMGINSYIYMDADTIPVD--K 371
Qy 407 LDVEYKALSSVNKTITVEYQRP--NENRTANI-QSMFTNIDTKNHT-----VSEQTIY 456
Db 372 KDVPFSVTIGNQITTTTADITYPAYKEADNNSIGSAFT--ETVSHVGNVEDPCGYNQVYV 429
Qy 457 INPLYSAKETNWNISNGDEGST-----IIDSSTIIKVKYKVDGNQNLPSNRIDY--SE 510
Db 430 VNPMDKDLGAKLKVAYHPKPTNIGQINQNVNTNIIKVRVPGYTL--NKGYDYNTND 486
Qy 511 YEDVTND--DYAQLGNNDVNFNGNIDSPYIIKVISKYDPNKDDVTTIQOVTMTOTTIN 568
Db 487 LVDVTDEFKRWYTGNSQVNLDFGDLTSAYVWVNTKFOYTNSSEPTLVQMATLSTGN 546
Qy 569 EYTGFEPTASYDNTIAFSTSSGQGGDLPPKPKYKIGDYWEDVDKDGIQ----- 618
Db 547 -----KSVSTGNALGFTNNQSGAG---QEVYKIGNVYWEDTNKNGVQELGKGVGVN 596
Qy 619 -----NTY----- 621
Db 597 TVTVFDNNTNTKVGAEAVTKEDGSLIPNLPNGDYRVEFNSLPKGYEVTPSKQGNBELDS 656
Qy 622 -----DN-----BKPNSNLVLT----- 634
Db 657 NGLSSVITVNGKDNLSADLGIYKPKNGLYVWEDTNKNGIQOQDEKSGISGVTVLKDEN 716
Qy 635 -----TYPDG----- 639
Db 717 GNVLKTVTTDADGKYFTDLNGYKVEFTTPGTYPTTTSQSDIEKDSNGLTITGVIN 776
Qy 640 -----TSKSV-----RT 646
Db 777 GADNMTLDSGFYKTPKYNLGNVYWEDTNKDGKQDSTEKGISGVTVTLKNGENGEVLQTTKT 836
Qy 647 DEBGKYQDGLKNGLTAKITFETPEGVYPTLKHSGTNPALDSEGS----- 692
Db 837 DKDGKYQFTGLENG-TYKVEFETPSGYTPTQVSGTDEGIDSGTSTGVIKDKNDTID 895
Qy 693 -----VW-----VTI-----NGQ----- 700
Db 896 SGFYKPTYNLGDVWEDTNKNGVQDKDEKISGVTVTLKDENDKVLKTVTTTDENGKYQFT 955
Qy 701 -----DMWTIDSGFYQTPKYS 716
Db 956 DLNNGTYKVEFETPSGYTPTSVTSGNTEKDSNGLTITGVIKADNNTLDSGFYKTPKYS 1015
Qy 717 LGNVYVWDTNKGIOGDDEKIGISGVKTLKDENGNIISITTTBNGKYQPDNLNSGYIV 776
Db 1016 LGDYVWYDSNKGQDSTEKIGIKDVKIVILLNEKEVIGITTKTDENGKYRFDNLDSGKYK 1075
Qy 777 HFKPSGMTQTTTDSGDDQDADGEVHVHTIITHDDFSDINGYDDE--SDSDSDSDSDS 835
Db 1076 IFKPTGLTQGTNTTDEDD-KDADGGEVDVTIITHDDFTLUNGYYEEETSDSDSDSDS 1134
Qy 836 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 895
```

```
Db 1135 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1194
Qy 896 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 955
Db 1195 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1254
Qy 956 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1015
Db 1255 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1309
Qy 1016 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1064
Db 1310 -----KSTKD-----KLPDTGANEDVGS- 1064
Qy 1065 KGTLLGTLTLAGLALL-GKRRNRNK 1089
Db 1360 NATLFGFLFAALGSLLLFGRRKKQNK 1385
RESULT 12
Q99W48 PRELIMINARY; PRT; 953 AA.
ID Q99W48;
DC Q99W48;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding
DE protein.
GN SDRC OR SAV0561 OR SA0519
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AF003359; BAB56723.1; -.
DR EMBL; AF003131; BAB41750.1; -.
DR PIR; C89824; C89824.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR008454; Cna B.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF05738; Cna B; 2.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG anchor; 1.
DR TIGRFAMs; TIGR01168; Ysirk signal; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Complete proteome.
SQ
SEQUENCE 953 AA; 103292 MW; 729A7169A074A1E5 CRC64;
Query Match 37.1%; Score 2093.5; DB 16; Length 953;
Best Local Similarity 43.8%; Pred. No. 3e-64;
Matches 488; Conservative 140; Mismatches 298; Indels 189; Gaps 28;
Qy 2 INKKNLLTKKPIANKSNKYAIRFTVGTASTIVIGATLLFGLGHNEAKAEENSVDQVKD 61
Db 1 MNKTKATNRKGMIPNRLNKFIRKYSVGTASLVGTLLFGLSGHEAKAAE----- 52
Qy 62 SNTDDELSDNSQSDSEKKNVINNNQSNINTDNNQIIKKEETNNYDGIKRSDEKTEST 121
```

```
Db 53 -HTNGELNQSNETTAPSEN-----KTT 74
Qy 122 TNVDENEATFLOKTPQDNTHLTREEVKESSESSVSSNSSIDTAQOPSHTTINREESVQTS 181
Db 75 EKVDSRQL-----KONTQTA-----TADQPKVT-----97
Qy 182 NVEDSHVDFPANGKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDEX---ISNQ 238
Db 98 -----MSDSATVKETSSNQSS-----PQNATASQSTTQTSNVTNDSSTYSNE 142
Qy 239 DELLNLPINEYENKARPLSTTSAQPSIKRVTVNQALAA-EQGSNVNHLIKVTQSI-----292
Db 143 TDKSNL--TQAKVSVTTPKTTIKRPNRMAVNTVAAPQOGTNNVDKVFHTNIDIAIDK 201
Qy 293 -----TEGVDDSGVIAKHAENLIYDVFTEVDKVKSGDTMTVDIDKNTVPSDLT 343
Db 201 GHVNTKNTGTEFWATSSDVUK-----LKANYTTIDDSVKEGDTFTFKYQGYFRGQSVR 252
Qy 344 DSFTIPKIDNSGEIATGYDNKNKQITFTDYVDKYENIKAKHLKLTYSIDKSKVPNN 403
Db 253 LPSQTNLYNAQGNIIAKGIYDSKNTTFTTYVDQYTNVSGSPQVAFAKRENATTD 312
Qy 404 NTKLDVEYKTALSSVNTIIVFYQRPENRNTANLQSMFTNIDTNKHTVEQTIYN-PLRY 462
Db 313 KTAYPEVTLGNDTYSKDVIVDY---GNQKGQOLISSTNYINNEDLSRNMNTVYVNPQKKT 369
Qy 463 SAKETNV-NISGNGDEGSIIDSTIIKVKYVGNQNLPSNRNIYDYSEVEDVTND-DYA 520
Db 370 YTKETFTVNTU-----GYKENPDAKPKIYEVDQNFQVDS-FTPDTSKLDKVGQFDVI 423
Qy 521 QLGNNDVNIINFGN-----IDSPYIIKVISYDPNKKDYYTIIQOTVTMTQTTINEYTGFEFT 576
Db 424 YSNDNKATATVLLNGQSSSDQYIIQQVAYPDNSSTDNGKIDYTLQNGKSSWSN---479
Qy 577 ASDNTIAFTSSGQCGDLPPEKTYKIGDYVEDVDKGIQNTDNKPLSNVLTLY 636
Db 480 -SYSNVNGSSTANGD-----QKYNLGDYVWEDTNKQKQDA--NEKGIVGVYILKD 529
Qy 637 PDGTS-KSVRTDGDQYKQDGLKNGLYKITFTPEGYTTLKHSNTGNPALDSEGNVWV 695
Db 530 SNGKELDRITTDENGKYQFTGLSNG-TYSVEFTSPAGYFTTANAGTDDAIVSDGLHTTG 588
Qy 696 TINGQDDMTIDSGFYOPPKYSLGNYVWYDNTKQGIQGDDEKGISGVKVTLKDENGNIIST 755
Db 589 VIKADNMWTLDSGFYKTPKYSGLDYVWYDSNKGQKQDSTTEKGIKGVKVTLQNEKGEVIGT 648
Qy 756 TTTDENGKYQFDNLNSGNYIVHFDKPSGMYTQTTTDSGDDDEQADGEEVHTITDHDPS 815
Db 649 TETDENGKYRFDNLDSGKYKVIPEKPAGLTQGTNTTEDD-KDADGGEVDVITTDHDDFT 707
Qy 816 IDNGYDDESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 875
Db 708 LDNGYEEET-SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 766
Qy 876 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 935
Db 767 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 826
Qy 936 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 995
Db 827 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 886
Qy 996 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1055
Db 887 DSDSDSADGKHTPKMSTV-----KQOHTAKALPE 918
Qy 1056 TGANEDYGSKGTLLGLTFLAGLALL-GKRKRNRK 1089
Db 919 TGSNNNSNNGTLFGGLFAALGSILLFGRKKQNK 953
```

```
Q8NXX7
ID Q8NXX7 PRELIMINARY; PRT; 955 AA.
AC Q8NXX7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ser-Asp rich fibrinogen-binding bone sialoprotein-binding
DE protein.
GN SDR OR MW0516.
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
DR EMBL; AP004823; BAB94381.1; -.
DR GO; GO:0016020; C-membrane; IEA.
DR InterPro; IPR008454; Cna_B.
DR InterPro; IPR005877; Gpos_Ysirk.
DR Pfam; PF05738; Cna_B; 2.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01167; LPXFG_anchor; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
KW Complete proteome.
SQ
SEQUENCE 955 AA; 103642 MW; 65FD8BB93A477A22 CRC64;

Query Match 36.8%; Score 2075.5; DB 16; Length 955;
Best Local Similarity 43.8%; Pred. No. 1.2e-63;
Matches 485; Conservative 144; Mismatches 307; Indels 171; Gaps 27;

Qy 2 INKNNLLKKPIANKSNKYAIRFTVGTASIVIGATLLFGLGHNEAKAEENSVDQVDK 61
Db 1 MNKKTVTRKGMIPNRLNKFIRKYSVGTAILVGLTILFGLSGHEAKAE-----52
Qy 62 SNTDDELSDNSQSSDEEKDYNVNNQSNINTDDNQIIEKEETNNYDIEKSEDETEST 121
Db 53 -HTNGELNQSNETTAPSEN-----KTT 74
Qy 122 TNVDENEATFLOKTPQDNTHLTREEVKESSESSVSSNSSIDTAQOPSHTTINREESVQTS 181
Db 75 EKVDSHQL-----KONTQTA-----TADQPKVT-----97
Qy 182 NVEDSHVDFPANGKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDEX---ISNQ 238
Db 98 -----MSDSATVKETSSNQSS-----PQNATASQSTTQTSNVTNDSSTYSNE 142
Qy 239 DELLNLPINEYENKARPLSTTSAQPSIKRVTVNQALAA-EQGSNVNHLIKVT--DOSITEG 295
Db 143 TDKSNL-TQAKDVSAATPKTTIKRPNRMAVNTVAAPQOGTNNVDKVFHSNIDIAIDK 201
Qy 296 Y---DDSEGVIAKHAENLIYDVFTEVDKVKSGDTMTVDIDKNTVPSDLTDSFTPIK 352
Db 202 HLNKOTGKTEFWATSSDVUKLKNANYTTIDDSVKEGDTFTFKYQGYFRGQSVRLPSQTNLY 261
Qy 353 DNSGEIATGYDNKNKQITFTDYVDKYENIKAKHLKLTYSIDKSKVPNNNTKLDVEYK 412
Db 262 NAQENLIAGIYDSTNTTFTTYVDQYTNVSGSEQVAFAKRENATTDKTAKEVS 321
Qy 413 TALSSVNTIIVFYQRPENRNTANLQSMFTNIDTNKHTVEQTIYN-PLRYSAKETNV-N 470
Db 322 LGNDTYSEIIVDY---GNKKAQPLISSTNYINNEDLSRNMNTAYVNPQKNTYTKQTFVN 378
Qy 471 ISGNGDEGSIIDSTIIKVKYVGNQNLPSNRNIYDYSEVEDVTNDYDQALGNNN---D 527
```

Db	379	LT-----GYKFNPAKFKIYVTDQNFQVDS-FTPDTSKLDKDVNTQNFNITYSDNKTAT	432	DR	TIGRFAM6; TIGR01167; LPXTG anchor; 1.
Qy	528	VNINFGNIDP---PYIIKIVSKDPKODYTTTQQVTMTTINEYTGEBFRFASYNITAF	585	DR	TIGRFAM5; TIGR01168; YSRK_signal; 1.
Db	433	VDLMNGQTSNKKQYIIQQVAYPNTSTDKGIDYTLDTDKTKYSWEN-----SYSNVNGS	487	KW	PROSITE; P50847; GRAM_POS_ANCHORING; 1.
Qy	586	STSSGOGQGLDPEKTKYKIGDYVWEDVDKDGQNTNDNEKPLSNVLATLTYPDGTS-KSV	644	SQ	SEQUENCE 1347 AA; 145958 MW; AB8D393884AD8034 CRC64;
Db	488	STANGD-----QKYNIGDYVWEDTNKDGQDA--NEKGIGGVVILKDSNGKELDR	538		
Qy	645	RTDEDKYQFDGLKGLTKYITPETPEGYTPTLKHSCTNPALDSEGNVWVVTINGODDMT	704		
Db	539	TTDENGKYQFTGLSNG-TYSEFSTAGYPTTANAGTDDAIVSDGLTGTGVKOADNMT	597		
Qy	705	ISGFFYQTPKYSIGNVYVWDTNKGIGQDDEKIGSVKVTLDKDNENIISTTTTDENGKY	764		
Db	598	LDSGFFYKTPKYSIGDYVWYDSNKGQDQSTEGIKGVKVTLDQNEKEVIGTTTDENGKY	657		
Qy	765	QPDNLNGYIVHFDKPSGWTQTTDSGDDDEQADGEEVHTITDHDDFSIDNGYYDDE	824		
Db	658	RFNLDGSKYKVIPEKPAQLTGNTTDD-KADGGEVDVTTDHDFTLNGYYEE	716		
Qy	825	-SDS	883		
Db	717	TSDDS	776		
Qy	884	DSDDS	943		
Db	777	DSDDS	836		
Qy	944	DSDDS	1003		
Db	837	DSDDS	896		
Qy	1004	DSDDS	1063		
Db	897	GKHTPAKPSVTV-----KQHKTAALPETGSENNNS	928		
Qy	1064	SKGTLGLTLPAGLGALL-GKRKNRK	1089		
Db	929	NNGTFGGLFAALGSLLLFGRRKKQNK	955		
RESULT 14					
Db	Q8NXX6	PRELIMINARY; PRT; 1347 AA.			
AC	Q8NXX6;				
DT	01-OCT-2002 (TrEMBLrel. 22, Created)				
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)				
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)				
DE	Ser-Asp rich fibrinogen-binding bone sialoprotein-binding				
DE	protein.				
GN	SDRD OR MW0517.				
OS	Staphylococcus aureus (strain MW2).				
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.				
OX	NCBI_TaxID=196620;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=2240717; PubMed=12044378;				
RA	Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,				
RA	Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,				
RA	Yamamoto K., Hiramatsu K.;				
RT	"Genome and virulence determinants of high virulence community-				
RT	acquired MRSA."				
RL	Lancet 359:1819-1827(2002).				
DR	EMBL; AP004823; BAB94382.1; -				
DR	GO; GO:0016020; C:membrane; IEA.				
DR	InterPro; IPR008454; Cna_B.				
DR	InterPro; IPR005877; Gpos_YSRK.				
DR	InterPro; IPR001899; Gram_pos_anchor.				
DR	Pfam; PF05738; Cna_B; 5.				
DR	Pfam; PF00746; Gram_pos_anchor; 1.				
DR	Pfam; PF04650; YSRK_signal; 1.				

Qy	1	MINKKN-NLLTKKKPIANKSNKAIARKFTVGPASIVIGATLFLGLGHNAKAENSQDV	59		
Db	1	MLARENKTAIRKGMVSNRLNKFIRKVTGTGTASILVGTTLFGLGNQAKAAESTNKL	60		
Qy	60	KDSNTDDELSDNSDSEBEKNDVINNNQSIINTDNNQIIKKEETNNYDGIKRSERDTE	119		
Db	61	NEATT-----SASDNQSSD--KVDMLQNLQEDNTKNDQ-----KEMVSSQGN	102		
Qy	120	STTNVDEN-EATFLQKTPQDNTHLTETEEVEKSSSVESNSSIDTAQOPSHTTINREESV	178		
Db	103	ITSNGKSKIEKESVQSTTGNKVEVSTAKSDEQSPKSTNEDLNTKQ-----TISQEA	157		
Qy	179	TSDNVEDSHVSDFANSKIKESNTESGKEBENTIEQPNKVKEDSTTSQPSGYTNIDEKIS	238		
Db	158	-PDLQENKSVVNAQPTNEENKVKDAKTESLTL---NVKSDAIKS-----NAETLVDN	206		
Qy	239	DELLMLPINEYENKARPLSTTSAQP-----SIKRVTVNQLAAEQGNSVNHILIKVTD	294		
Db	207	SNSNN-----ENNADIILPKSTAPKRLNTRMRIAAPSPSTEAKNVNDLITSTTLTV	260		
Qy	295	GYDSEGVKKAHAENLIVDTFEVDDDKVSGDWTVDIDKNTV-----PSDLTSFTI	348		
Db	261	DADKNKIVPAQDYELKSKI--KVDDDKVSGDYFTIKY-SDTVQVYGLNLPDRIK--I	314		
Qy	349	PKIKD-NSGEIATATYDNKKNKQITTYTFTDYVDKYENIKAHILKLSIYIDSKVPNNNTKL	407		
Db	315	GDIKDPNGETATAKHDTANLLIYTFDYVDRENSVQMGINSYIMDADTIP--VSKN	372		
Qy	408	DVEYKALSSVNNKTITVEYQRPN--ENRTANLQSMFTNIDTKNHT-----VEQTII	457		
Db	373	DVEFNVTIGNTTKTNTANIQYDPYVNEKNSIGSAFT--ETVSHVGNKENPGYKQTIYI	430		
Qy	458	NPLRYSAKETVNI-----SGNGDEGSTIIDDSTIIKVKVGNQNLPSDNLRYDYS--EY	511		
Db	431	NPSENSLNNAKLVQAYHSSYPNNIQQINKEVTDIKIYQVPGYTL---NKGVDNTEKEL	487		
Qy	512	EDVTNDYQAQ---LGNNDNINFGNIDSPYIIKVISYDPNKKDDYTTTQQVTVMOTTIN	568		
Db	488	TDVTN-QYLQKITYGDNNSAVIDFGNADSAVVMVNTKRFQYTTSESPTLVQVNTLS---	542		
Qy	569	EYTGEBFRFASYNITAFSTSSGQGDLPPEKTKYKIGDYVWEDVDKDGIGQ-----	618		
Db	543	--SDNSKSASMGNALGFTNNQSGAG---QSVYKIGNYVWEDTNKNGVQELGEKGVGNV	596		
Qy	619	-----NTN-----	621		
Db	597	TVTVFDNNTNKVGEAVTKEDGSYILPNLPNGDYRVEFSPNLKPGYEVTPSKQGNNEELS	656		
Qy	622	-----DN-----EKPLSNVLVTI---	634		
Db	657	NGLSSVITVNGKDNLSADLGIYKPKYNLGDYVWEDTNKNGIQDQDEKIGSVTVTLKDN	716		
Qy	635	-----TPDG-----	639		
Db	717	GNVLKVTVTDADGKYKFTDLNGNYKVEFTTPEGVPTTVTSGSDIEKDSNGLTGTGIN	776		
Qy	640	-----TSKSV-----	646		
Db	777	GADNMTLDSGFYKTPKYNLGNVYWBNTKDGKQDSTERKIGISGVTVTLKNENGEVLQTKT	836		
Qy	647	DEDKYQFDGLKGLTKYITPETPEGYTPTLKHSCTNPALDSEGN	692		
Db	837	DKDGKYQFTGLENG-TYKVEFETPSGYTPTQVSGTDEGIDSNGTSTTGVIKDKNDTID	895		



Db 826 SDS 885  
Qy 1001 SDSDSDSDSVSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1060  
Db 886 SDAGKHTPAKPMSTV-----KQHKAKALPETGSEN 917  
Qy 1061 DYGSKGTLGTLFAGIGALL-LGKREKNRK 1089  
Db 918 NNSNNGTLFGGLFAALGSLLSFGRRKKQNK 947

Search completed: October 5, 2004, 20:52:22  
Job time : 143.638 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:34:15 ; Search time 22.8315 Seconds  
(without alignments)  
2490.441 Million cell updates/sec

Title: US-10-806-288-15  
Perfect score: 5646  
Sequence: 1 MINKKNLLTKKKPIANKN.....FAGLALLKKRKNKNKN 1092

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1319.5	23.4	1637	1 MRSP_STAAU	P80544 staphylococ
2	708.5	12.5	1253	1 DPPP_HUMAN	Q9nzv4 homo sapien
3	675.5	12.0	934	1 DPPP_MOUSE	P97399 mus musculu
4	606.5	10.7	687	1 DPPP_RAT	Q62598 rattus norv
5	554	9.8	406	1 SR40_YEAST	P32583 saccharomyc
6	487.5	8.6	1018	1 FNBA_STAAU	P14738 staphylococ
7	471	8.3	640	1 SANT_PLAFA	P13821 plasmodium
8	388.5	6.9	695	1 ARP_EUGGR	Q04732 euglena gra
9	368	6.5	1664	1 SLPI_CLOTM	Q06852 clostridium
10	361.5	6.4	400	1 RTOA_DICDI	P54681 dictyostell
11	343.5	6.1	1658	1 YM67_YEAST	Q03661 saccharomyc
12	335.5	5.9	4910	1 MDN1_YEAST	Q12019 saccharomyc
13	333	5.9	797	1 VG48_HSVSA	Q10103 herpesvirus
14	331.5	5.9	1419	1 ALAI_CANAL	O13368 candida alb
15	315	5.6	937	1 HYL_CANAL	P46591 candida alb
16	314	5.6	2492	1 ATRX_PANTR	Q7ygm4 pan troglod
17	311	5.5	2452	1 RPB1_PLAFA	P14248 plasmodium
18	310	5.5	1189	1 YJH6_YEAST	P47035 saccharomyc
19	303	5.4	2492	1 ATRX_HUMAN	P46100 homo sapien
20	301.5	5.3	503	1 DMPI_MOUSE	O55188 mus musculu
21	297	5.3	2492	1 ATRX_PONPY	Q7ygm3 pongo pygma
22	295.5	5.2	489	1 DMPI_RAT	P98193 rattus norv
23	294.5	5.2	429	1 DR48_YEAST	P18899 saccharomyc
24	291	5.2	1744	1 TANA_XENLA	Q01550 xenopus lae
25	288	5.1	389	1 SERI_BOMMO	P07856 bombyx mori
26	285.5	5.1	513	1 DMPI_HUMAN	Q13316 homo sapien
27	283.5	5.0	1337	1 DEXT_STRDO	P39693 screpococc
28	283	5.0	914	1 PBPA_BACSU	P39793 bacillus su
29	276.5	4.9	1085	1 IFH1_YEAST	P39520 saccharomyc
30	276	4.9	500	1 GAR2_SCHPO	P41891 schizosacch
31	271.5	4.8	1025	1 MK21_YEAST	Q12176 saccharomyc
32	271	4.8	688	1 LIP_STAEP	Q02510 staphylococ
33	271	4.8	2476	1 ATRX_MOUSE	Q61687 mus musculu

ALIGNMENTS

RESULT 1

ID	MRSP_STAAU	STANDARD;	PRT;	1637 AA.
AC	P80544; Q9ZF62;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Methicillin-resistant surface protein precursor.			
GN	PUS.			
OS	Staphylococcus aureus.			
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.			
OX	NCBI_TaxID=1280;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Isolate 1061;			
RX	MEDLINE=21189215; PubMed=11292719;			
RA	Savolainen K., Paulin L., Westerlund-Wikstrom B., Foster T.J.,			
RA	Korhonen T.K., Kuusela P.;			
RT	"Expression of pls, a gene closely associated with the mecA gene of			
RT	methicillin-resistant Staphylococcus aureus, prevents bacterial			
RT	adhesion in vitro.";			
RL	Infect. Immun. 69:3013-3020(2001).			
RN	[2]			
RP	SEQUENCE OF 624-628; 676-682; 938-948; 1156-1168; 1176-1185;			
RP	1199-1205 AND 1217-1224.			
RC	STRAIN=Isolate 1061;			
RX	MEDLINE=96270743; PubMed=8665912;			
RA	Hilden P., Savolainen K., Tynnelae J., Vuento M., Kuusela P.;			
RT	"Purification and characterization of a plasmin-sensitive surface			
RT	protein of Staphylococcus aureus.";			
RL	Eur. J. Biochem. 236:904-910(1996).			
CC	-I- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by			
CC	an amide bond (Potential).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AF115379; AAD09131.1; -			
DR	InterPro; IPR005877; Gpos_Y5IRK.			
DR	InterPro; IPR001899; Gram_pos_anchor.			
DR	Pfam; PF00746; Gram_pos_anchor; 1.			
DR	Pfam; PF04650; Y5IRK signal; 1.			
DR	TIGRFAMS; TIGR01167; LPXTG anchor; 1.			
DR	TIGRFAMS; TIGR01168; Y5IRK signal; 1.			
DR	PROSITE; PS00847; GRAM_POS_ANCHORING; 1.			
KW	Antibiotic resistance; Glycoprotein; Cell wall; Peptidoglycan-anchor;			
FT	Repeat; Signal.			
FT	SIGNAL 1 48			
FT	CHAIN 49 1601			
FT	PROPEP 1602 1637			
FT	DOMAIN 1301 1582			
FT	POTENTIAL.			
FT	METHICILLIN-RESISTANT SURFACE PROTEIN.			
FT	REMOVED BY SORTASE (POTENTIAL).			
FT	141 X 2 AA TANDEM REPEATS OF D-[SAG].			

FT SITE 1598 1602 LPXTG SORTING SIGNAL (POTENTIAL).  
 MOD\_RES 1601 1601 AMIDE-LINKED TO CELL WALL (POTENTIAL).  
 SQ SEQUENCE 1637 AA; 174573 MW; 75BE9ADB469BD309 CWC64;

Query Match  
 Best Local Similarity 23.4%; Score 1319.5; DB 1; Length 1637;  
 Matches 461; Conservative 180; Mismatches 401; Indels 635; Gaps 60;

QY 7 NLLTKKXP--JANKSNKYAIRKFTVGTASIVIGATLLFLGLHNEAKABENSVDQKDSNT 64  
 DB 2 NKNSKKLDFLPNKLKNSIRFTVGTASILVGTALIFGVANDQAEAAENNTTKQDSS 61  
 QY 65 D-----DELSDNDQSSDEEKDVNNOSINTDNNQ 97  
 DB 62 DASKVGNVQTEOSSANSNESDIPQVDVTKTTEQASTEKEANTTEQASTEKADTTE 121  
 QY 98 IIKKEETNNYDGIK-----RSEDRTESTNVNDEATFLQ 133  
 DB 122 QATTEAPKAGTDKVTETAPKAEETDKATTEAPKAEETDKATTEAPKAEETDKATTE 181  
 QY 134 KTP---QNTHLTEEVKSSVSSNSSIDTAQPSHTTINREESVQTSNV----- 183  
 DB 182 EAPAAEETSKAETAPKAEETSKAETAPKAEETSKAETAPKAEETDKVETEEAPK 241  
 QY 184 --EDSHVSDFANSKIKESNTESGKEENTIEQPNKVKEDST-----TSQPS--G 227  
 DB 242 AEETSKAATEKAPKAEETKVTETAPKAEETKVTETAPKAEETKVTETAPKAEETK 301  
 QY 228 YTNIDEKISNOELLNLPINEYE---NKAPPLSTTS-AQPSIKRVTVNOLAASOGSNVN 282  
 DB 302 RTQVVDVTA-KDLYKKESEVTAETAEKVLKPLKDISLSNEEIKKIALSEVLKETANKEN 360  
 QY 283 HLIKVDOSITEGVDDSGVKAHAENLIYDVT-----FEVDDKVKSGD---TMTVIDIK 335  
 DB 361 AQRATERSVSNART-----NVNYSATALRAAAQDVTVKKGTFNGTAHGDIIH 410  
 QY 336 NTVPSDLTDSFTPIKIDKNSGEIATGT-----YDNK---NKQIYTT----- 374  
 DB 411 KTYKEEFPNEGTLTAFNTFNP--NTGKGGALEYNDKIDFNKOFITVPVANNQNGTTG 468  
 QY 375 -----FT-----DYVDKYENIKAHKLKLTYSIDKSKVPNNNTKLDVEYKLTALSSVNKTI 422  
 DB 469 ADGWGFMFTQGNQDFLNQGGILR-----DKGMANASGFKIDTAYNNVNGKVDKL- 518  
 QY 423 TVEQRNENRANLQSM-----FTN-----IDTKNHTVEQTIYIN- 458  
 DB 519 -----DADKTNLSQIGAAKVGYGTGVKNGADGVTVNQGNALNTKDKPVNKKIYADN 571  
 QY 459 -----PLRYSARETNV----- 469  
 DB 572 TTNHLDGQFGHQRNDVVLNDAATSTITATYAGTKWKATDDLGIDKSKQYNPLITSSH 631  
 QY 470 --NISNGND-----EGSTI-----IDDSITIK----- 492  
 DB 632 MONRYSNGIMRTNLEGVITTPQADLDDVEVTKQPIPHKTIREFDPLEPGSPDVIQK 691  
 QY 493 VGDQNQL-----PSNRIYDSE-YEDVNDQYAL-----GNNNDVINNF- 532  
 DB 692 GEDGEKTTTPTKVDPTGDVVERGEPTTEVTKNPVDVIVHFTPEEVPQGHKDFDNL 751  
 QY 533 -----GNIDSPYIKVISKYDNKDQYT----- 555  
 DB 752 IDGTEEVPKPGIKNKPETGEVVT---PPVDDVTKHGPKAGEPVTKEEIPFEKKREPNPD 808  
 QY 556 -----TIQQTVMQTIN----- 568  
 DB 809 LKPGEEKVTQBGQGEKTTTPTTINPLTGEKVGEGEPTTEVTKEPVEIDITQFGGEVVP 868  
 QY 569 -----EY-----TGFRFASVDNTIAFSTSSGQCG- 593  
 DB 869 GHKDFDNLPIDGTREEVPKPGIKNKPETGEVVTTPVDDVTKHGPKAGEPVTKEEIPFE 928  
 QY 594 -----GDLPP-----EKT-----YKIG-----DYVWEDVD----- 613

## RESULT 2

DSPP\_HUMAN STANDARD; PRT; 1253 AA.  
 ID DSPP\_HUMAN  
 AC Q9NZW4; O95815;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Dentin sialoprophosphoprotein precursor [Contains: Dentin phosphoprotein  
 (Dentin phosphophoryn) (DPP); Dentin sialoprotein (DSP)].  
 GN DSPP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20168992; PubMed=10706475;  
 RA Gu K., Chang S.R., Ritchie H.H., Clarkson B.H., Rutherford R.B.;  
 RT "Molecular cloning of a human dentin sialoprophosphoprotein gene.";  
 RL Eur. J. Oral Sci. 108:35-42(2000).  
 RN [2]  
 RP SEQUENCE OF 463-1253 FROM N.A.  
 RC TISSUE=Tooth;  
 RX MEDLINE=99094526; PubMed=9879917;  
 RA Gu K., Chang S.R., Slaven M.S., Clarkson B.H., Rutherford R.B.,





```
QY 148 K-ESSSVESNSSIDTAQPSHTTIRRE-----SVQTSNVDERS-HVSDPANSKIKESWT 201
DB 154 HGQAGQNTKSGGASDVSONGATLVQNEPPEASIKNSTNHEAGIRGSGVATHTTTPQRE 213
QY 202 ESCKEENTLEOPNKVEDS-----TTSQPSY-TNIDEKISNQDELLNPLINEYENKARPL 256
DB 214 GLGSENGQTEVTPSIGEDAGLDTDGSPNGVGEDEDTGSGD-----256
QY 257 STTSAQPSIKRV-----TVNQLAAEQSNVNHLIKVTQDSITEGYDPS-----EGVIKAHDAE 309
DB 257 -GEGAAGDGRSHDGTGQGGQSGHGNTDHR---GQSSVSTEDDSSKEQEGPNGHNGD 312
QY 310 NLIYDVTFFVDVKVSGDTWTVIDIKNTV-PSLTD-----STPIKIKONSSEITA 360
DB 313 N-----SSEENGVEEGDSTQATQDKELKSPKDRDAEGGIIISQSBACPSGKSQDQGIET 366
QY 361 TGYDNK-NKOITVYTFDYVVKYENIKAKHLKLTYSIDKSKVPNNNTKLDEYKATLSSVN 419
DB 367 EG--PNKGKNSII-----TKESGKLSGSKDSNGHQGVLDKNS-----403
QY 420 KTIIVYQRPENRTANLQSMFTNIDTKNHTVBQTIYINPLRYSAKETNVTNISNGDEGS 479
DB 404 -----PKQGES-----DKPQGTAEK-----SAAHNLGHSRIGSSN 435
QY 480 TIIDSTIIKVKGNQNLPSNRNYDYSEYEDVTNDDYALGNNDVNIINFGNTDSPY 539
DB 436 SDGHDS-----YFDDDESQMDPKSDSDESGDSDT-----NSESANESGSRGDASY 484
QY 540 IIKVISKYDPNKDDYTTIQQVTVMQTTINEYTFRTASVDNTIAFTSGGQGGQLPPE 599
DB 485 TSDESDDDNDSDH-----AGEDD--SDDSGGDSDS-----516
QY 600 KTYKIGDYVWEDVDKQGIQNT-NDNEKPLSNVLTLTPDGTGSKSVRTD--EDGKYQFDG 656
DB 517 ---SNGDGSSEDKDSDSDHDSN-----DSEKSDSDSDSDSDSDSDSDSDS 561
QY 657 LKNGLYKTFETPEGYTPTLKHSGTNPALDSGNSVWVTINGQDDMTIDSGYQTPKY$ 716
DB 562 -----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 604
QY 717 LGNVVYDT-----NKDGIQDDEKISGVKVKTLKBNGNIISTTTDENGKYQFN 768
DB 605 -----DTCDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 657
QY 769 LNSGNYIVHFKPSGMTQTITDSCGDDQDADGEEVHVITTDHDDFSDNGYVYDDESDD 828
DB 658 SDSDS-----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 709
QY 829 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 888
DB 710 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 768
QY 889 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 948
DB 769 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 828
QY 949 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1008
DB 829 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 888
QY 1009 SDSDSDSDSDSDSDSDG---SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1051
DB 889 SKDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 933
```

## RESULT 4

DSDP RAT

ID DSDP RAT STANDARD; PRT; 687 AA.

AC Q62598; P70578; Q9R057;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

```
DE Dentin sialophosphoprotein precursor (Contains: Dentin phosphoprotein
DE (Dentin phosphophoryn) (DPP); Dentin sialoprotein (DSP)).
GN DSDP OR RDSPP2.
OC Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM DPP-2).
RC STRAIN=Sprague-Dawley;
RC MEDLINE=20435277; PubMed=10978503;
RA Ritchie H.H., Wang L.-H.;
RT "The presence of multiple rat DSP-pp transcripts.";
RL Biochim. Biophys. Acta 1493:27-32(2000).
RN [2]
RP SEQUENCE OF 1-387 FROM N.A., AND SEQUENCE OF 18-26.
RC STRAIN=Sprague-Dawley; TISSUE=Odontoblast;
RX MEDLINE=94148875; PubMed=8106414;
RA Ritchie H.H., Hou H., Veis A., Butler W.T.;
RT "Cloning and sequence determination of rat dentin sialoprotein, a
RT novel dentin protein.";
RL J. Biol. Chem. 269:3698-3702(1994).
RN [3]
RP SEQUENCE OF 421-687 FROM N.A. (ISOFORM DPP-1).
RC STRAIN=Sprague-Dawley;
RX MEDLINE=96355551; PubMed=8702961;
RA Ritchie H.H., Wang L.-H.;
RT "Sequence determination of an extremely acidic rat dentin
RT phosphoprotein.";
RL J. Biol. Chem. 271:21695-21698(1996).
RN [4]
RP SEQUENCE OF 1-19 FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=99333695; PubMed=10403786;
RA Yamazaki H., Kunisada T., Miyamoto A., Tagawa H., Hayashi S.-I.;
RT "Tooth-specific expression conferred by the regulatory sequences of
RT rat dentin sialoprotein gene in transgenic mice";
RL Biochem. Biophys. Res. Commun. 260:433-440(1999).
RN [5]
RP SEQUENCE OF 29-33; 70-79; 93-109; 136-148; 162-188; 266-308; 398-423
RP AND 426-438, AND PHOSPHORYLATION OF SER-292 AND SER-298.
RX MEDLINE=21125612; PubMed=11042175;
RA Qin C., Cook R.G., Orkiszewski R.S., Butler W.T.;
RT "Identification and characterization of the carboxyl-terminal region
RT of rat dentin sialoprotein.";
RL J. Biol. Chem. 276:904-909(2001).
RN [6]
RP TISSUE SPECIFICITY.
RX MEDLINE=98055479; PubMed=9395101;
RA Ritchie H.H., Berry J.E., Somerman M.J., Hanks C.T., Bronckers A.L.,
RA Hottot D., Papagerakis P., Berdal A., Butler W.T.;
RT "Dentin sialoprotein (DSP) transcripts: developmentally-sustained
RT expression in odontoblasts and transient expression in
RT pre-ameloblasts.";
RL Eur. J. Oral Sci. 105:405-413(1997).
CC -!- FUNCTION: DSP may be an important factor in dentinogenesis. DPP
CC may bind high amount of calcium and facilitate initial
CC mineralization of dentin matrix collagen as well as regulate the
CC size and shape of the crystals.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- ALTERNATIVE PRODUCTS.
CC Event=Alternative splicing; Named isoforms=2;
CC Name=DPP-1; Synonyms=PP240;
CC IsoId=Q62598-1; Sequence=Displayed;
CC Name=DPP-2; Synonyms=PP171;
CC IsoId=Q62598-2; Sequence=VSP_003855;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Specifically expressed in teeth, mainly in
CC odontoblasts and transiently in pre-ameloblasts.
CC -!- PTM: DSP is glycosylated.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to a
CC frameshift in position 380.
CC -----
```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AF247187; AAK96895.1; --  
CC EMBL; U02074; AAA18932.1; ALT\_FRAME.  
CC EMBL; U63111; AAC52774.1; --  
CC EMBL; AF114987; AA048588.1; ALT\_TERM.  
KW Bimernization; Extracellular matrix; Signal; Glycoprotein;  
KW Phosphorylation; Sialic acid; Alternative splicing.  
FT SIGNAL 1 17  
FT CHAIN 18 687 DENTIN SIALOPHOSPHOPROTEIN.  
FT CHAIN 18 447 DENTIN SIALOPHOSPHOPROTEIN.  
FT CHAIN 448 687 DENTIN SIALOPHOSPHOPROTEIN.  
FT MOD\_RES 57 57 PHOSPHORYLATION (BY CK2) (POTENTIAL).  
FT MOD\_RES 226 226 PHOSPHORYLATION (BY CK2) (POTENTIAL).  
FT MOD\_RES 253 253 PHOSPHORYLATION (BY CK1) (POTENTIAL).  
FT MOD\_RES 278 278 PHOSPHORYLATION (BY CK1) (POTENTIAL).  
FT MOD\_RES 292 292 PHOSPHORYLATION (BY CK2).  
FT MOD\_RES 298 298 PHOSPHORYLATION (BY CK2).  
FT MOD\_RES 315 315 PHOSPHORYLATION (BY CK2) (POTENTIAL).  
FT MOD\_RES 319 319 PHOSPHORYLATION (BY CK2) (POTENTIAL).  
FT MOD\_RES 329 329 PHOSPHORYLATION (BY CK2) (POTENTIAL).  
FT MOD\_RES 337 337 PHOSPHORYLATION (BY CK2) (POTENTIAL).  
FT MOD\_RES 345 345 PHOSPHORYLATION (BY CK2) (POTENTIAL).  
FT MOD\_RES 366 366 PHOSPHORYLATION (BY CK1) (POTENTIAL).  
FT CARBOHYD 55 55 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 82 82 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 128 128 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 312 312 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 369 369 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT VARSPLIC 567 635 Missing (in isoform DPP-2).  
FT FTID=VSP\_003855.  
FT CONFLICT 74 74 N -> D (IN REF. 5; AA SEQUENCE).  
FT CONFLICT 564 564 S -> T (IN REF. 3).  
SQ SEQUENCE 687 AA; 70179 MW; 9A845EED6AA31B63 CRC64;  
Query Match 10.78; Score 606.5; DB 1; Length 687;  
Best Local Similarity 32.08; Pred. No. 1.4e-16;  
Matches 211; Conservative 68; Mismatches 250; Indels 131; Gaps 30;  
453 QTVINPLRYSAKET-----NWNISGNGEGST-IIDSTIIKVKY-GDN 496  
97 QPILANAQNSAKEREVETVYHGDHAGGENTANGIRQGVIAENAEKESKVGQP 156  
497 QNLPDSNRIYDYGEYDVT---NDVYALQNNNDVNFNIDSPYIIKISKYDPNKOD 553  
157 HQDTKTGLASDTQNGDALVQENEPQVAGSKNSTHEVTHGSGV---AAQETTPQREG 213  
554 YTIQVTVMTVINEYTGFRASVDNTAFSTSSGQGGDLPPEKTYKIGYVWEDV- 612  
214 EGENQGAETVPSIGEGAG-----LDNT-----EGSPSGNGIBEDEDTGS-GDVGADAG 262  
613 -----DKDGIQ-----NTNDNEKPLSNVLVTLPDGTGSKSVRTDEDKYQFDGLKNG 660  
263 DGRESHDTGTEHGCGSGGNNDR-----CQGSVSTEDDDSKQEGSPNG 307  
661 LTYKITPETPEGYPTPLKHSCTNPALD-----SEGNSVWVYTINGQDDMTIDSGFYTPKY 715  
308 -----RGCDNTSSSEETGIEBGDGTQTTQDNQLNSPTEGGIISQAEA 349  
716 SLGNVYVYDNTKDIGIQQDDEKGLS-GVKVTLKDNENIISTTTDENGKY--QFDNLNSG 772  
350 C-----PSGQSONQLETEGSGTGNKSSITKESGKL--SGSKDNGHGMELDKRNSP 400  
773 NYIVHFDKPSGMTQ---TTTDSGD---DEQDADGEEVHTIITDHDFFSIDNGYYDDES 826  
401 KQ-GESEKPGQAAEKSDTHNMGMHSRIGSSNSDGHDSY-----DFDESQOG---DDPNS 452

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; L11275; AAA35091.1; --  
CC EMBL; X73541; CAA51946.1; --  
CC EMBL; Z28317; CAA82171.1; --  
CC PIR; S38170; S38170.  
CC Germonline; 140071; --  
CC SGD; S0001800; SRP40.  
CC GO; GO:0005730; C:nucleolus; IDA.  
CC InterPro; IPR007718; SRP40\_C.  
CC Pfam; PF05022; SRP40\_C; 1.  
CC DOMAIN 25 374 ASP/SER-RICH  
CC CONFLICT 400 400 G -> N (IN REF. 1).  
CC SEQUENCE 406 AA; 41015 MW; 8EA007695AF4BA1D CRC64;

Query Match 9.8%; Score 554; DB 1; Length 406;  
Best Local Similarity 46.8%; Pred. No. 7.5e-15;  
Matches 137; Conservative 25; Mismatches 93; Indels 38; Gaps 6;

QY 827 SDSDSDSDSDSDSDS-----DSDSDSDSDSDSDSDSDSDSDSDSDS 874  
DB 453 SDESNGSGSDDDANSATENGHGDAVTSDESSNGSDSHAGEDSSDDTSDTD-D 511  
QY 875 SDSDSDSDSDS-DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 933  
DB 512 SDNSGDDSDSEKDKDESDNSNHDNDSDSEKSDS-SDSDSDSDSDSDSDSSETSDS 570  
QY 934 DS 992  
DB 571 SDSSDTSDS 629  
QY 993 SDSDSDSDSDSDSDS-DSVSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1051  
DB 630 DSDSDSSSDS 686

RESULT 5  
SR40 YEAST  
ID SR40 YEAST STANDARD; PRT; 406 AA.  
AC P32583;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Suppressor protein SRP40.  
GN SRP40 OR YKR092C OR YKR412A.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RC STRAIN=ATCC 28383 / FL100;  
RA Lalo D., Carles C., Sentenac A., Thuriaux P.;  
RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94205265; PubMed=8154186;  
RA Bou G., Esteban P.F., Baladron V., Gonzalez G.A., Cantalejo J.G.,  
RA Remacha M., Jimenez A., del Rey F., Ballesta J.P.G., Revuelta J.L.;  
RT The complete sequence of a 15,820 bp segment of Saccharomyces  
RT cerevisiae chromosome XI contains the UBI2 and MPL1 genes and three  
RT new open reading frames."  
RL Yeast 9:1349-1354(1993).  
CC -!- FUNCTION: Not known; weak suppressor of a mutant of the  
CC subunit AC40 of DNA dependant RNA polymerase I and III.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).











```
Best Local Similarity 21.1%, Pred. No. 4.3e-07;
Matches 243; Conservative 163; Mismatches 425; Indels 320; Gaps 56;

QY 167 SHTTINRE--ESVQSDNVEDS---HVSDFANSKIKESNTE--SGKEENTIEQPNKVKED 219
D 19 STTSVNNSPAEATPSIEWLDTKEVHVGVITATIKVNNIRKLAGVQLN-----IKFD 71
QY 220 STTSQPSGYNIDKISQDELNLNLPINEY---ENKARPLSTTSQAQSIKRVTV----- 270
D 72 PEVLQP-----VDPATGEEPTDKSMFVNRVLLTNSKYGP--TPVAGNDIKSGIINFATGY 124
QY 271 NQLAAEQGSNNVH-----LIKVDQSITGEYDDSD---EGVKA-----HDAENL 311
D 125 NNTAYKSSGIDEHTGIGIKFVKLKKQNTSIR--FEDTSLMFGAISGTSFLFDWDAETI 182
QY 312 I-YD-----VTFEVD-DKVKSGD--TMTVDID--KNTVPSDLT---D 344
D 183 TGEVIOQDLIVVEAEPLKDAVALELDKTKVKVGDIITATIKIENKNFAGVQLNIKYD 242
QY 345 SFTIPKIKDNGSIIATGTY-----DNKNK-----QITYTPTDYVDKYENIKAH 388
D 243 PTMLEALELGTGSAIAKRTWPVTGGTVLQSDNYGKTTAVANDVGAGIINFARAYSNLTKY 302
QY 389 LKL-----TSYIDK---SKVPNNNTKLDVEYKTA--SSVNTK----- 421
D 303 RETGVAEETGIGIKGFRVLKAGSTAIRPEDTTAMPGAIEGTMYMFDWYGENIKGYSVVQP 362
QY 422 --ITVEYQRNENRNTANLQSMFTNIDPKNHTVQTI-----YI----- 457
D 363 GEIVAEGEETPEPTPEPTVPTVPTVTEPVPSELPDSYVIMELDKTKVKVGDII 422
QY 458 -----NPLRYSAKETNNVNSNGDESGSTIIDSTIHK-VYKVGDNQNLNPDNRIYDYS 509
D 423 TATIKIENKNFAGVQLNIKYDPTMLEAIELETSIAKRTWPVTGGTVLQSDN---YG 478
QY 510 EYEDVTNDYDQAOLGNNDVNNIGNDSPYIIKVISKYDNKDDYTIQOTVTMOTINE 569
D 479 KTTAVANDVGAGI-----INFAYSN-----LTKY-----RETGVAE 511
QY 570 YTG-----EPTASDYNT-IAFSTSGQGGDLPP--EKTYKIGDYVWEDVDXGI--- 617
D 512 ETGLIGIKGRVLKAGSTAIRFEDTTA---NPGALEGTYMF-DWGENIKGYSVVQP 565
QY 618 QNTNDNKKPLSNVLVLT---YPDGTSKSVRTDEGKY-----Q 653
D 566 EIVAEGEETPEPTPEPTVPTVPTVTEPVPSELPDSYVIMELDKTKVKEGDVIIATIR 625
QY 654 FDGLKGLTKITFEITPEGTPTLKHGNTPALDSEGNVWVITNG-----QDDM----- 703
D 626 VNNIKNLAGYQIGIK-----YDPKVLFAFNIEGTGDIPEGTPAVGGTILKNRDYLTGVA 691
QY 704 --TIDSGFYQTPKYSGLNYYVYDTNKGDIQDDDEK--GISGVKVTILKDENGNI-----IST 755
D 682 INNVSIGILNFAAY-----YVYFDYREEGKSEDTGIIGNGFV-LKAEDTTRFEELES 736
QY 756 TTTDENGKYQFD---NLNNGNYIVH-----FDKPSGMWQ--TTTSGDD 794
D 737 MPGSIDGTYMLDYLNRISGVYVYIQAPAKASDEPIPTDTPSDEPTPSDEPTSP 796
QY 795 DEQADAGEEHHVITDHDHDSINGVYDDESDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 846
D 797 SDEPTPSDEPTPEPTPEPTPTTP--SDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPT 855
QY 847 -----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 894
D 856 EEPITDTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPEPTPTDTPSDEPTPSD 915
QY 895 SDSDSDSDSDSDSDSDSDSDSD-----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 947
D 916 EPTPSDEPTPSDEPTPSDEPTPEPTPEPTPTDTPSDEPTPSDEPTPSDEPTPSDEPTPS 975
QY 948 DSDSDSDSDSDSDSDSDSD-----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1000
```

```
Db 976 DBPTPSDEPTPSDEPTPSDEPTPEPTPTDTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTP 1035
QY 1001 SPSDSDSDSDSVSDS-----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSKDKLP- 1054
D 1036 SDEPTPSDEPTPSDEPTPEPTPTDTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPT 1095
QY 1055 -----DTGANE 1060
D 1096 EPIPTDTPSDE 1106

RESULT 10
RTOA_DICDI
ID RTOA_DICDI STANDARD; PRT; 400 AA.
AC P54681;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Protein rtoA (Ratio-A).
GN RTOA.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97108762; PubMed=8951083;
RA Wood S.A., Ammann R.R., Brock D.A., Li L., Spann T., Gomer R.H.;
RT "rtoA links initial cell type choice to the cell cycle in
RL Dictyostelium";
RL Development 122:3677-3685(1996).
CC -!- FUNCTION: May have dual functions, one in vegetative cells at
CC starvation involving cell-type choice and a later function
CC involving progression to the tipped mound stage.
CC -!- DEVELOPMENTAL STAGE: Expressed in vegetative cells. Levels
CC decrease dramatically as development begins, and remain low
CC throughout the aggregation and loose mound stages. The levels
CC become more abundant again at the tight-mound stage and remain
CC high throughout the rest of development and fruiting body
CC formation.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U48298; AAC47311.1; -.
CC DictyBase; DB016851; rtoA.
CC Transmembrane; Repeat.
CC TRANSMEM 30 50
CC DOMAIN 30 38
CC DOMAIN 87 96
CC DOMAIN 181 181
CC REPEAT 181 190
CC REPEAT 191 200
CC REPEAT 201 211
CC REPEAT 212 222
CC REPEAT 223 233
CC REPEAT 234 244
CC REPEAT 245 255
CC REPEAT 256 266
CC REPEAT 267 277
CC REPEAT 278 287
CC REPEAT 288 298
CC SEQUENCE 400 AA; 39840 MW; 0392F4E68DC27A75 CRC64;
Query Match 6.4%; Score 361.5; DB 1; Length 400;
Best Local Similarity 30.1%; Pred. No. 1.5e-07;
Matches 100; Conservative 48; Mismatches 123; Indels 61; Gaps 5;
```

```
QY 744 TLKDEGNIIITTTTDENGKYQFDNLNSGNYIVHFDKPKSCMTQTITDSDGDDDDQADAGE 803
Db 49 TIKKNGSIGSS-----SQSLSEVDSISSG-- 77
QY 804 VHVITDHDHDFSDNGYDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 863
Db 78 -----SNETASSEGVSSSSNSGSGSTNSGSEAGSSNSGSGSTNSGSEA 124
QY 864 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 923
Db 125 SGSSNSGSGSGSTDSNSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 183
QY 924 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 983
Db 184 NSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 242
QY 984 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1043
Db 243 SGSSNSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 302
QY 1044 SSKST-----KDKLPDTGANE 1060
Db 303 SSGKTKTCSFHDTLSTINTVDDDEIECTGKE 334

RESULT 11
Ym67 YEAST STANDARD; PRT; 1658 AA.
AC Q03661; Q04988;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical 187.1 kDa protein in GUAL-ERG8 intergenic region.
GN YMR219W OR YMR261.13 OR YM9959.01.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313268; PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jags R., Lyle G., Moulé S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT XIII."
RL Nature 387:90-93 (1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z49809; CAA89934.1; -
DR EMBL; Z49939; CAA90190.1; -
DR PIR; S55101; S55101.
DR Germonline; 142894; -.
DR SGD; S0004832; ESC1.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0006348; P:chromatin silencing at telomere; IMP.
KW Hypothetical protein.
SQ SEQUENCE 1658 AA; 187137 MW; 3893F968305A757D CRC64;

Query Match 6.1%; Score 343.5; DB 1; Length 1658;
Best Local Similarity 19.2%; Pred. No. 3.6e-06;
Matches 241; Conservative 215; Mismatches 446; Indels 355; Gaps 57;

QY 47 NEAKAENSVDVKDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 101
```

```
Db 232 DEBYAEEGALQDV-----SNDEYAEBSGQVERKN--IGQEQA-NVENATQISSS 277
QY 102 E--ETNMY-DGIEKRSRSTSTTNVDENEATFLQKTPQ-----DNT----- 140
Db 278 DSSEGQNYSEGVEMELEDDIDVESDAEKESQCAEGTEHSVDFSKYMQPRTDNTKIPVIE 337
QY 141 -----HLTEBEVKSSSVE--SSNSIIDTAQOPSHITI-----NREESVQTS 180
Db 338 KYESDEHKVHYRSEDGAFDFGSVNIISVDSDEDEBSQAESYASANAENYVHNEHEDDK 397
QY 181 DNVEDSHVDFANSKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYT--NIDEKISNQ 238
Db 398 ELIEDIESDSESQAQES--EQGSB--DFEYKMKNEKSTSEETENTSESRDQQA-K 451
QY 239 DELLALPINEYENKARPLSTTSAQPSI-KRVTVNQLAAEQGSNNVHLIKVTQSITE--- 294
Db 452 DAYTKNKVQEQENDEPEKDDIIRSSLDKNFHNKSEYSENV--LENETDPAIVEREN 509
QY 295 GYDDSEGVKAHAENLIYDVTFEVDVKVSGDTMTVDIDKNTVPSDLTDSFTPIKIDN 354
Db 510 QINDVEG-----YDVT-----GKSVESDLHHS-PDNLVLAARAMLQFQ 548
QY 355 SGELIATGYDNKNKQITVTFTDYVD-----KYENIKAKHLKLTYSIDKSKVNNNTKL 407
Db 549 QSR---NSNCPQKEQVESYLGHSNGSNLSRSLDSESEIQPLKDFTCEN---NNLKT 602
QY 408 DVEYKTALSSVNKTIIVEYQRPENRTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKET 467
Db 603 D-----RGDLSSVEIEVEKSEKLDG-----STEKELV-----PL---STD 638
QY 468 NVNISNGDEGTI--IDSTIIKVKYKGDQNQLPD-----SNRIYDVS 509
Db 639 TINNSLGNEDSIYSLDDADAI-----SENLTDPVLEIKTTPKYEVWISSEYVST 691
QY 510 EYEDVT-----NDYQAQLGNNDVNIN-----FCNIDSPVI 540
Db 692 SHEDNTVAMPQVEYTPPMNDPFPNSLNDDEYK---KHDLKSLTAAAPAFKKADEFV 748
QY 541 IKVISK-----YDPNKD-----DYTTIQOTVTMTTINEYTGFEFTASVDNTI 583
Db 749 EAGVTKSLTSTSGHTNIFHTSKETKQVSLDDESTENVTFE---NENTGDNKQSKNPP 805
QY 584 AFSTSGGQGGDLPRKTYKIGDYWEDVDKGIQNTNDNEKPLSNVLTLTPD---GT 640
Db 806 GVANSTDKSTEDNTDEKYFSAINY-----TNVTGSSCEDIIETASNVLENRLYCEKDMNE 861
QY 641 SKSVRTDEGKYQFDGLKNGLTAKITFETPEGYTPTLKHSGTNPALDSGNSVWVTINGQ 700
Db 862 AEMSGDECVKQNDGSKTQISF--STDSPDNF-----QESNDN--TEFSSTKYKVRNS 911
QY 701 DDMTIDSGFYQTPKYSGLNYVWYDYNKGIQ---GDDEKIGISGVKVTLKDE--NGNIIST 755
Db 912 DLEDDSLAKELTAEVVDKLEDESEDSYEQDYADPEPG-----NDEGSNENIVKG 963
QY 756 TTTDENGKYQFON-----LNSGNVIVHFDKPKSCMTQTITDSDGDDD 795
Db 964 TKKOTLGIYEPEENKYNKVHBEETLFEANVSSVNVQNKDMETDVINQEAQANYEAGERK 1023
QY 796 E--QDADGEEVHTIDH--DDFSIDN-----818
Db 1024 YIQNTDTEEAHISIIERIDENAIQNNMIEPERSVCVEKTHNEVLFERRATTIENTVALEN 1083
QY 819 --GYDDESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 873
Db 1084 NTNMHDQVQACSDSDRDQDSTAERNKVEGSAKHNLDIRVSSSEIESEVEPLKPSDRSNIF 1143
QY 874 DS-----DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 903
Db 1144 SSPIRVIGAVKGVKGVDAESFVKKIDVMDSESDNDVIGDYNQDIFNKSNS--TDASV 1202
QY 904 DSDSDSDSDSDSDSD-----SDSDSDSDSDSD-----DSDSDSDSDSDSDSDSDSDSD 949
```



```
Db 4005 RYNNAMTVLERSRIYYTNTSRGMYILSTILSLAKNGFCSPQPPSEEVDDKNLQEG--- 4061
Qy 712 TPKYSLGNYVYDTNKGIGQDEKSGVVKVTLKDENGNIISTTTTDENGKYQFNLS 771
Db 4062 ---TGLGDGSGCAQNNKDVQDEED-----LTEDAQNE-----KEQDKDERDENE 4105
Qy 772 GNYIVHFKPSGWTQTITDSGDDDEQADAGEEVHV-----TITDHDDFSIDNGYYDDES 826
Db 4106 DDAVEGMDWAGELEDLSNGEENDEEDTSEBELDEEIDDLNEDDPNAILDDQWDDKA- 4164
Qy 827 SDSDSDSDSDSDD-SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 885
Db 4165 SDNSKEKTDONLDGKNQEDVQAENDEQQRNKEGGEDPNAPEDGEEIENDENAE 4224
Qy 886 DSOSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 944
Db 4225 ENDVGEQDEVKDEEGE-DLEANVPETETLDLPEDMMLDSEHESEDVDMSCMPDOLN 4283
Qy 945 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1004
Db 4284 KEEVGNDEEVKQESGIESNENDEPGFEEDAGETETALDEERGABEDVMTNDEKCE 4343
Qy 1005 SDSDSDSVSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1064
Db 4344 ENGPEQAMSDERELKQDAAMENKEKGGQNTGLDGVEEKADTEDIDQEAQVQDSGS 4403
Qy 1065 KG 1066
Db 4404 KG 4405

RESULT 13
VG48 HSVSA STANDARD; PRT; 797 AA.
AC Q01033;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Hypothetical gene 48 protein.
GN 48 OR EDLFS.
OS Herpesvirus saimiri (strain 11).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI_TaxID=10383;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9233368; PubMed=1321287;
RA Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Biesinger B.,
RA Newman C., Wittmann S., Craxton M.A., Coleman H., Fleckenstein B.,
RA Honess R.W.;
RT "Primary structure of the herpesvirus saimiri genome.";
RL J. Virol. 66:5047-5058(1992).
CC -I- SIMILARITY: TO EBV BRP2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X64346; CAA45671.1; -
CC InterPro; IPR008550; DUF832.
CC Pfam; PF05734; DUF832; 1.
KW Hypothetical protein.
SQ SEQUENCE 797 AA; 89867 MW; 9E294234AD850E23 CRC64;

Query Match 5.9%; Score 333; DB 1; Length 797;
Best Local Similarity 18.8%; Pred. No. 3.9e-06;
Matches 116; Conservative 104; Mismatches 275; Indels 122; Gaps 22;

Qy 506 YDSEYEDVTNDYQAQGNNDNVNFGNIDSPYIIKVISKDPNKDDYTTIQVTMQT 565
```

```
Db 194 YDFNEHK-----NEISLVATCINCCWLFMLQY--MSDDLAEALN-KT 237
Qy 566 TINEYTGFEPTASYDNTIAFTSSGQGG-----GDLPPKTYKIGYVWEDVDKGI 617
Db 238 YLALHPND--KASYSNLIKFLTNSHREHVTKVNVKAFMQSSLYKI-----I 283
Qy 618 QNTNDNEKPLSNLV-----TLTYP--DGTSKSVRTDEDGKYQFDGL 657
Db 284 KDTKNPSPKTKMLMISILGSRIGMDLFCOSVLKAPLIDHKLSVPSEYED----FD-- 337
Qy 658 KNGLTVKIFETPEGYPTLKHSGTNPALDSGNSV-----WVTINGQD----- 701
Db 338 EBEVELCISDDEV-----SEDGNLCVLDDESESVNSVALRQVLTVDQANEKEYKKII 391
Qy 702 DMTIPSGFYQTPKYSIGNVYVYTNKGIQGDDEKIGISGVKVTLDKENGNIISTTTT 761
Db 392 DKSDDRDRDRDKDEYLENEE--YNRDEEDEDGEDEKDEKEEGEDGDD--GEDEGEDE 448
Qy 762 GKYQFNLNSGNYIVHFKPSGWTQTITDSGDDDEQADAGEEVHV-VTITDHDDFSIDNGY 820
Db 449 GEDEGEDEG-----DEGEDEGEDEDEDEGEDEGEDEGEDEGEDEG 491
Qy 821 YDESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 880
Db 492 EGDEGEDEG-EGDEGKDEGDEGKDEGDEGDEGDEGDEGDEGDEGDEGDEGDEG 550
Qy 881 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 940
Db 551 EDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEG 610
Qy 941 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1000
Db 611 EGDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDE 670
Qy 1001 SDSDSDSDSVSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1060
Db 671 GDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDE 725
Qy 1061 -----DYGS---KGTLL 1068
Db 726 VVQNPFYVYNLQKSTL 742

RESULT 14
ALAI CANAL STANDARD; PRT; 1419 AA.
ID ALAI CANAL
AC O13368;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Agglutinin-like protein ALAI precursor (Agglutinin-like adhesin).
GN ALAI OR ALG5.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98053977; PubMed=9393828;
RA Gaur N.K., Klotz S.A.;
RT "Expression, cloning, and characterization of a Candida albicans
RT gene, ALAI, that confers adherence properties upon Saccharomyces
RT cerevisiae for extracellular matrix proteins.";
RL Infect. Immun. 65:5289-5294(1997).
CC -I- FUNCTION: MAY PLAY A ROLE IN ADHESION AND PATHOGENESIS.
CC -I- PTM: N-glycosylated and O-glycosylated (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC
```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```
CC DR EMBL; AF025429; AAB88883.1; --
CC DR PIR; T30531; T30531; Candida ALS.
CC DR InterPro; IPR008440; Candida ALS.
CC DR Pfam; PF05792; Candida ALS; 1.
CC KW Cell adhesion; Glycoprotein; Repeat; signal.
FT SIGNAL 1 17
FT CHAIN 18 1419 AGGLUTININ-LIKE PROTEIN ALA1.
FT DOMAIN 399 404 POLY-THR.
FT DOMAIN 408 418 POLY-THR.
FT DOMAIN 437 441 POLY-THR.
FT DOMAIN 673 676 POLY-SER.
FT DOMAIN 687 690 POLY-SER.
FT DOMAIN 700 703 POLY-SER.
FT DOMAIN 719 724 POLY-SER.
FT DOMAIN 749 752 POLY-SER.
FT DOMAIN 787 791 POLY-SER.
FT DOMAIN 869 872 POLY-SER.
FT DOMAIN 875 883 POLY-SER.
FT DOMAIN 901 911 POLY-SER.
FT DOMAIN 1216 1221 POLY-SER.
FT CARBOHYD 665 665 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 919 919 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1301 1301 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1326 1326 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1419 AA; 149635 MW; 249F33F688A9D5B6 CRC64;

Query Match 5.9%; Score 331.5; DB 1; Length 1419;
Best Local Similarity 23.0%; Pred. No. 8.7e-06;
Matches 237; Conservative 116; Mismatches 366; Indels 313; Gaps 43;

QY 130 TFLQKTP-----QDNTHLTEREKSSVESNSSIDTAQOPSHTTNREES----- 176
DB 66 TFLNMECVKFTASQKSVLDTAGVXYATCFISGEFTTFSLKCTVNNLRSSIKAL 125
QY 177 ----VQTSNVEDSHVS-DFANSKIKESNTES-----GKEENTIEQPNKVKEDSTTSQPS 226
DB 126 GTVTLPFAFNVGGTSGSSVDLEDSKCFAGTNTVTFNDGSKKLSI-----AVNPEKSTVDQS 181
QY 227 GYTNIDEKINSQDELLNLPINEYENKARPLSTTSAQPSIKRVTVNQLA--AEQGSNVNHL 284
DB 182 GY-----LTSRFMPLNKLATLYVAPOCNGYT----- 210
QY 285 IKVTDQSITEGYDDSEGVIKAHDAENL-----YDVTTFEVDKVKSGDTMTVDI 333
DB 211 -----SGTWGFTSYGDV-AIDCSNVHIGISGVNDNHPVTSEFSYTKSCSFGISI 263
QY 334 DKNTVPS---DLTDSFTPIKIKONGSEIIATGYDNNKQKITTYFTDYV---DKYENIKA 387
DB 264 TYQNVAGYRPFIDAYISPS-----DNNQYQLSYK-NDYTCVDDYQWQHAPP 308
QY 388 HLKLTSYIDSKVPNN-----NTKLDVEYKTALS-----SVNKTITVEYQRPENRTAN 436
DB 309 TLKWTGYKNDAGSNGIIVATRTVTDSTVATVTLFPNPSVDKTKTIELQIP---TTT 366
QY 437 LQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVTNINISGNGDEGSTIIDSTIIKRVYKVGON 496
DB 367 ITTSYVGVTTIS-----YSTKTAPI-----GETATVIVD----- 394
QY 497 QNLPSNRIDYSEYEDVNDVDAQLGNNDNINFGNIDSPYIIKVIKIDPNKDDYTT 556
DB 395 -----VPYHTTTVTSEWTGTTITTTTTRNPTDSIDT-VVQVPS---ENPATTITT 441
QY 557 --IQQVTMTQTNNE-----YTGFEFTASVDNTIAPSTSSGQGGDL 596
DB 442 QFWSESTSTTTIYNLSKGTDSVIVREPHTVTTTTFWSESFATTETIYSK----- 493
QY 597 PPEKTYKIGYVWEDVDKQIGNTNDNEKPLSNLV-----TLTYPDGTSKSVRTDEPG 650
DB 494 -PEGT-----DSVIVREPHTVTTTTFWSESFATTET- 525

RESULT 15
HYR1 CANAL
ID HYR1_CANAL STANDARD; PRT; 937 AA.
AC P46591;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Hyphally regulated protein precursor.
GN HYR1.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]_
SEQUENCE FROM N.A.
RC STRAIN=ATCC 10261;
RX MEDLINE=96404781; PubMed=8808922;
RA Bailey D.A., Feldmann P.J.F., Bovey M., Gow N.A.R., Brown A.J.P.;
RT "The Candida albicans HYR1 gene, which is activated in response to
RT hyphal development, belongs to a gene family encoding yeast cell wall
RT proteins."
RL J. Bacteriol. 178:5353-5360 (1996).
CC -!- FUNCTION: Nonsential component of the hyphal cell wall
CC (Potential).
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (Potential).
CC -!- DEVELOPMENTAL STAGE: Abundant in hyphae.
CC -!- INDUCTION: Induced specifically in response to hyphal development.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:42:37 ; Search time 39.7921 Seconds  
(without alignments)  
2639.748 Million cell updates/sec

Title: US-10-806-288-15  
Perfect score: 5646  
Sequence: 1 MINKKNLLTKKPIANKSN.....FAGLGALLLKRRKRNKN 1092

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: Pirl:\*  
2: Pirl2:\*  
3: Pirl3:\*  
4: Pirl4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5646	100.0	1092	2 T30214	fibrinogen-binding
2	2698.5	47.8	1166	2 T28680	fibrinogen-binding
3	2589	45.9	1141	2 E89824	hypothetical prote
4	2234	39.6	1385	2 D89824	hypothetical prote
5	2031.5	37.1	953	2 C89824	hypothetical prote
6	1880	33.3	1315	2 T28679	fibrinogen-binding
7	1652	29.3	989	2 D89852	fibrinogen-binding
8	1628	28.8	933	2 S41539	fibrinogen-binding
9	1536	27.2	877	2 F90070	Clumping factor B
10	785	13.9	882	2 AG1671	probable peptidogl
11	782.5	13.9	903	2 AG1299	probable peptidogl
12	624.5	11.1	2271	2 F90073	hypothetical prote
13	554	9.8	406	2 S38170	SRP40 protein - ye
14	554	9.8	3394	2 T18501	hypothetical prote
15	549	9.7	2570	2 T17451	fimbriae-associate
16	525.5	9.3	1038	2 H90053	hypothetical prote
17	520	9.2	940	2 S19702	fibronectin-bindin
18	502.5	8.9	961	2 G90053	hypothetical prote
19	487.5	8.6	1018	2 A32192	fibronectin-bindin
20	471	8.3	334	2 A54138	acidic repetitive
21	471	8.3	640	2 A54502	S antigen precursor
22	452	8.0	2910	2 T28156	DNA-directed RNA p
23	442.5	7.8	1999	2 AB2018	hypothetical prote
24	435	7.7	1192	2 A71623	probable secreted
25	430.5	7.6	4550	2 T18440	hypothetical prote
26	391.5	6.9	1217	2 S52714	sericin1B - silkw
27	388.5	6.9	695	2 S27390	calcium-binding pr
28	381.5	6.8	955	2 T18435	hypothetical prote
29	379.5	6.7	3724	2 T18427	hypothetical prote

30	371.5	6.6	1305	2 T00670	probable inositol
31	370	6.6	1063	2 D86731	hypothetical prote
32	368	6.5	1664	2 T18262	S-layer protein -
33	366	6.5	3844	2 T18402	asparagine/asparta
34	364.5	6.5	792	2 T42963	hypothetical prote
35	363.5	6.4	1360	2 T18403	asparagine/asparta
36	362	6.4	1193	2 G71605	hypothetical prote
37	351.5	6.2	1428	2 T08852	lustrin A - Califo
38	345.5	6.1	2481	2 D90011	FmcB protein (lmpo
39	343.5	6.1	1658	2 S55101	hypothetical prote
40	338	6.0	4776	2 E95206	cell wall surface
41	336	6.0	2485	1 H71621	serine/threonine-s
42	335.5	5.9	4910	2 S64942	probable membrane
43	333.5	5.9	1072	2 A86827	hypothetical prote
44	333	5.9	797	2 A36811	hypothetical prote
45	331.5	5.9	1419	2 T30531	agglutinin-like ad

ALIGNMENTS

RESULT 1

T30214  
fibrinogen-binding protein - Staphylococcus epidermidis  
C:Species: Staphylococcus epidermidis  
C;Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 11-May-2000  
C;Accession: T30214  
R;Nilsson, M.; Frykberg, L.; Flock, J.I.; Pei, L.; Lindberg, M.; Guss, B.  
Infect. Immun. 66, 2666-2673, 1998  
A;Title: A Fibrinogen-binding protein of Staphylococcus epidermidis.  
A;Reference number: Z20781; MUID:98261511; PMID:9596732  
A;Accession: T30214  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-1092 <N1>  
A;Cross-references: EMBL:Y17116; NID:e1296734; PID:e1296735; PIDN:CAA76638.1

Query Match 100.0%; Score 5646; DB 2; Length 1092;  
Best Local Similarity 100.0%; Pred. No. 3.3e-211;  
Matches 1092; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MINKKNLLTKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNKAAEENSVDVK	60
DB	1	MINKKNLLTKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNKAAEENSVDVK	60
QY	61	DSNTDDELSDNDQSSDEEKNDVINNNQSIINTDDNNQIIKKEETNNYDGIKSESDRTES	120
DB	61	DSNTDDELSDNDQSSDEEKNDVINNNQSIINTDDNNQIIKKEETNNYDGIKSESDRTES	120
QY	121	TTNVDENEATFLOKTPQDNTHLTTEEVKSSSVSSSSIDTAQPSHTTINREESVQTS	180
DB	121	TTNVDENEATFLOKTPQDNTHLTTEEVKSSSVSSSSIDTAQPSHTTINREESVQTS	180
QY	181	DNVEDSHVSDPANSKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQDE	240
DB	181	DNVEDSHVSDPANSKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQDE	240
QY	241	LLNLPINEYENKARPLSTTSAQPSIKRVTNVQLAAEQGSNNVHLIKVTDQSIIEGYDDSE	300
DB	241	LLNLPINEYENKARPLSTTSAQPSIKRVTNVQLAAEQGSNNVHLIKVTDQSIIEGYDDSE	300
QY	301	GVTKAHDAAENLIYDVTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFTPIKIKNSGEIIA	360
DB	301	GVTKAHDAAENLIYDVTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFTPIKIKNSGEIIA	360
QY	361	TGYDNNKQIITFTDYDVKYENIKAHKLTSYIDKSKVPNNNTKLDVEYKTKALSNNK	420
DB	361	TGYDNNKQIITFTDYDVKYENIKAHKLTSYIDKSKVPNNNTKLDVEYKTKALSNNK	420
QY	421	TTITVEYQRPNNENRTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVNSGNDGSGT	480
DB	421	TTITVEYQRPNNENRTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVNSGNDGSGT	480

```
QY 481 IIDDSTIIKVKVGNQNLPSNRNIYDYSEVEDVTNDYALQGNNDVNIINFGNIDSPYI 540
Db 481 IIDDSTIIKVKVGNQNLPSNRNIYDYSEVEDVTNDYALQGNNDVNIINFGNIDSPYI 540
QY 541 IKVLSKIDPNKDDYTTIOQVTMTTINEXYTGERTASYDNTIAFSTSSGQGGDLPEK 600
Db 541 IKVLSKIDPNKDDYTTIOQVTMTTINEXYTGERTASYDNTIAFSTSSGQGGDLPEK 600
QY 601 TYKIGDYVWEDVDKDGIOQNTDNKPLSNVLVLTITPDGTSKSVRTDEDKYQFDGLKNG 660
Db 601 TYKIGDYVWEDVDKDGIOQNTDNKPLSNVLVLTITPDGTSKSVRTDEDKYQFDGLKNG 660
QY 661 LTYKIFTEPGYPTLKHSGTNPALDSEGNVWVTINGQDDMTIDSGFYQTPKYSIGNY 720
Db 661 LTYKIFTEPGYPTLKHSGTNPALDSEGNVWVTINGQDDMTIDSGFYQTPKYSIGNY 720
QY 721 VNYDTNKDGIQDDDEKIGSGVKVTLKDENGNIISTTTTDENGKYQFDNLASGNVIVHFDK 780
Db 721 VNYDTNKDGIQDDDEKIGSGVKVTLKDENGNIISTTTTDENGKYQFDNLASGNVIVHFDK 780
QY 781 PSMGTQTTTDSGDDDEQADGEEVHVITITDHDDFSIDNGYDDSDSDSDSDSDSDSDSD 840
Db 781 PSMGTQTTTDSGDDDEQADGEEVHVITITDHDDFSIDNGYDDSDSDSDSDSDSDSDSD 840
QY 841 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 900
Db 841 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 900
QY 901 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 960
Db 901 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 960
QY 961 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1020
Db 961 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1020
QY 1021 SDGSGSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1080
Db 1021 SDGSGSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1080
QY 1081 LGKRRKNRKNKN 1092
Db 1081 LGKRRKNRKNKN 1092

RESULT 2
T28680
fibrinogen-binding protein homolog - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-May-2000
C:Accession: T28680
R:Josefsson, E.; McCrea, K.; Ni Eidhin, D.; O'Connell, D.; Cox, J.; Hook, M.; Foster, T.
Microbiology 144, 3387-3395, 1998
A:Title: Three new members of the serine-aspartate repeat protein multigene family of Staphylococcus aureus
A:Reference number: 220510; MUID:99098700; PMID:9884231
A:Accession: T28680
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1166 <JOS>
A:Cross-references: EMBL:AJ005647; NID:e1318793; PID:e1318794; PIDN:CAA06652.1
A:Genetics:
A:Gene: sdE

Query Match 47.8%; Score 2698.5; DB 2; Length 1166;
Best Local Similarity 48.2%; Pred. No. 3.3e-97;
Matches 595; Conservative 158; Mismatches 267; Indels 215; Gaps 26;

QY 1 MINKN-NLTKKPTANKSNKAIKPTVGTASIVIGATLLPGLGHNKAKAENSVDV 59
Db 1 MINKN-NLTKKPTANKSNKAIKPTVGTASIVIGATLLPGLGHNKAKAENSVDV 59
QY 60 KDSNTDDELSDNQSSDEEKNDVNNNQSIINTDDNNQIIKKEETNNYDGIKRSDRTE 119
Db 60 KDSNTDDELSDNQSSDEEKNDVNNNQSIINTDDNNQIIKKEETNNYDGIKRSDRTE 119
```

```
Db 59 ENAKQDDATTSNDKNEVSVETENNSTNNSTNP-----IKKE--TNTDSQPEAKKESTS 110
QY 120 STTNVDENEATFLQKTPQDNTHLTETEEVKES---SSVESNSSSIDTAQOPSHHTINREES 176
Db 111 SSTQKQNNVATATETKPN--LEKENVAPSTDKTATEDITSVLEKEKAPNNT--ANNVDT 166
QY 177 VQTSNDNVEDSHVDFANSKIKESNT---BSGKEENTIEQPNKVKEDSTTSQPSGYTNIDE 233
Db 167 TKFS-----TSEPSTSEIQTKPTTPQESTNIENSQOPQTPSKVD---NQVTDATNPKE 216
QY 234 KIS-NODELLNLP-----INEVENKARPLUSTTSAOPSIKRV-----TVNQLAABQ 277
Db 217 PNVVSEELKNPEKELKELVRNDSNTDHSKPVATAPTSVAPKRVAKRFAVAQPAAVA 276
QY 278 GSNVNLHILKVTQOSITIEGYDDSGVILKAHAENLIYDVTFEVDDKVKSGDGTMTVDTKNT 337
Db 277 SNNVDOLILKVTQITIKVG--DGKNVAAAHGDKDIEYDTEFTIDNKVKKGDTMTINVDKNV 335
QY 338 VPSDLTDSFTPIPKIKNSGEIATGTYDNKNKQITTYFTDYVDKYENIKAHKLKLTSYIDK 397
Db 336 IPSDLTDKNPIDIITDPSGEVIAKGFDKATKQITTYFTDYVDKYEDIKSRLTLYSIDK 395
QY 398 SKYPANNTKLDVEYKTALSSVANKTITVEYQRPENRNTANLQSMFTNIDTNHTVBTIYI 457
Db 396 KTVV-NETSLNLTFAATAGKETSQNVTVDYQDPWVHGDSNIQSIFTKLBDEKQTIQOIYV 454
QY 458 NPLRYSAKETNVNINSGN-----DEGSTIIDSTIIKVKVGNQNLPSNRNIYDY 508
Db 455 NPLKASATWTKVDIAGSQVDDYGNIKLNGSIIIDQNTIEIKVKVNSDQQLPQSNRIYDF 514
QY 509 SEYEDVTND-DYALQGNNDVNIINFGNIDSPYIIVISKYDPNKKDDYTTIOQVTMTT 567
Db 515 SQYEDVTSQDNKKSFSNNVATLDFGDISAVIIVKVSQYTPTSDEGELDIAQGTSWRTT- 573
QY 568 NEYTGEPRTASVDNTIAPSTSSGQGGDL-PEPKYKIGDYVWEDVDKGIQNTDNEKP 626
Db 574 DKY-GYINTAGYSNFIIVTSNDTGGGDTGTVKPEKLYKIGDYVWEDVDKGVQGTSGKEP 632
QY 627 LSNVLVLTLYPDGTSKSVRTDSDGKYQFQGLKNGLTYYKITFTPEGTPTLKHSGTNPAL 686
Db 633 MANVLVLTLYPDGTSKSVRTDANGHYEFGGLKDGTYTVKFTFGYLPFKVNGITDGEK 692
QY 687 DSEGNVWVTINGQDDMTIDSGFYQTPKYSIGNYVWYDNTNKGIOGDDSKGISGVKVTULK 746
Db 693 DSGSSVTVKINGKDDMSLDTGFKYKPNGLDYVWEDTNDKGIQDANEPGIKDKVKTULK 752
QY 747 DENGNIIISTTT-----
Db 753 DSTGKVIIGITTTTDSAGKYKFTDLNNGNYVTFETPAGYTPVKNTTADDKDSNGLTTTG 812
QY 758 -----
Db 813 IKDADNMTLDRGFYKTPKYSGLGVWYDSNKGDKQDSTEKIGIKDVTVTI-LQNEKEVIGTT 872
QY 758 -TDENGKYQFDNLASGNVIVHFDKPSGVTQTTTSDGDDDEQADGEEVHVITITDHDPSI 816
Db 873 KTDENGKYRFDNLDSGKYKVIPEKPAGLTQVTNTTEDD-KDADGGEVDVITITDHDFTL 931
QY 817 DNGYVDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 876
Db 932 DNGYFEEDT-----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 980
QY 877 SPDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 936
Db 981 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1040
QY 937 SPDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 996
Db 1041 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1100
QY 997 SPDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSKDK-LPD 1055
Db 1101 SDSDSDAGKHTPVKPM-----TTKDHINKAKALPE 1131
```

[illegible]





[illegible]





Db 409 YIKGYQDK-----IESSGKVSATDTKRIIFEVNDTSKLSDSYYADPNDSNLKVTDFKN 464  
QY 504 RIYDYSEYEDVTNDYDIAQLGNNDVNINFGNIDSPYIIKIVSKYDPNKKDDVTT--IOQTV 561  
Db 465 RIY-----HPNVASIKFGDITKTYVVLVEGHYDNTGKKLKTQVIOENV 509  
QY 562 TMOQTINYEFTASVDNTIAFSTSSGGQGGDLPEPKYKIGDYVWEDVDKGIQNTN 621  
Db 510-----DPVTNRDYSI--FGW-----NN 524  
QY 622 DNEKPLSNVLVLTYPDGTSKSVRTDEGKYQFGLKNGLYTKITFTPEGYTPTLKXSG 681  
Db 525 EN-----VRYGGGSA-----DG-----DSAVN-----PKDPTP----- 548  
QY 682 TNPALDSEGSVWVTINGQDDMTIDSGFYQPKYSLGNYVYDNTKDGICQDDEKGISGV 741  
Db 549 -GPPVDPESP-----DPE----- 561  
QY 742 KVTLDKENGNIISTTTDENGKYQFNLNSGNIVYHFDKPSGMTQTTTSDGDDDEQADG 801  
Db 562-----PEPTPDPPEPSPDPEPSPDP 584  
QY 802 EEVHTITDHDPSIDNGYDDESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 861  
Db 585-----DSGSDSGSDSGSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 624  
QY 862 DS 921  
Db 625 DS 684  
QY 922 DS 981  
Db 685 DS 744  
QY 982 DSD 1041  
Db 745 DSD 804  
QY 1042 G-----NSSDKSTK-----DKLPDTGANEDYSGKGTLLGTLFAGLGAL 1079  
Db 805 RVTPPNNEQKAPSNPKGEVNHNSKVSQKHTDALPETG-DKSENTNATLFGAMWALLGSL 863  
QY 1080 LL-CKRKNKKNK 1091  
Db 864 LLFRKRQKHKEK 876

RESULT 10  
AG1671  
Probable peptidoglycan bound protein (LPXTG motif) lln1913 [imported] - Listeria innocua  
C;Species: Listeria innocua  
C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C;Accession: AG1671  
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, A.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: AG1671  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-882 <GLA>  
A;Cross-references: GB:AL592022; PIDN:CAC97143.1; PID:gl6414414; GSPDB:GN00178  
A;Experimental source: strain Clp11262  
C;Genetics:  
A;Gene: lln1913

Query Match 13.9%; Score 785; DB 2; Length 882;  
Best Local Similarity 35.1%; Pred. No. 2.2e-23;  
Matches 151; Conservative 130; Mismatches 91; Indels 58; Gaps 9;

QY 662 TYKITFETPEGYTPTLKXSGT-----NPAIDSEGSVWVTINGQDDMTI 705  
Db 95 TYPYFELSELSTIILSNPNIKANTKIDYKIAYLGIKGNLGNQ-----TVNGSS----- 145  
QY 706 DSGFYQPKYSLGNYVYDNTKDGICQDDEKGISGVK-----VTLKDENGNI 752  
Db 146 NNFFIDTSRNAIGARV---NHLGV-----GVSSVSTFTLTIDLLALGVLTALPSANDGK 196  
QY 753 ISITTTDENGKYQFNLNSGNIVYHFDKPSGMTQTTTSDG-DDEQADGEEVHVITDH 811  
Db 197 LDFARTGDLGLVDLLNSN-----AARGFITTDVGADADADADADADADA 247  
QY 812 DPFSDNGYVYDDESDS 871  
Db 248 DA 307  
QY 872 DS 931  
Db 308 DA 367  
QY 932 DS 991  
Db 368 DA 427  
QY 992 DSD 1051  
Db 428 DAD 486  
QY 1052 KLPTGANED 1061  
Db 487 ADADADADAD 496

RESULT 11  
AG1299  
Probable peptidoglycan bound protein (LPXTG motif) lmo1799 [imported] - Listeria monocytogenes  
C;Species: Listeria monocytogenes  
C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C;Accession: AG1299  
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, A.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: AG1299  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-903 <GLA>  
A;Cross-references: GB:NC 003210; PIDN:CAC99877.1; PID:gl6411253; GSPDB:GN00177  
A;Experimental source: strain EGD-e  
C;Genetics:  
A;Gene: lmo1799

Query Match 13.9%; Score 782.5; DB 2; Length 903;  
Best Local Similarity 35.4%; Pred. No. 2.8e-23;  
Matches 154; Conservative 131; Mismatches 83; Indels 67; Gaps 10;

QY 662 TYKITFETPEGYTPTLKXSGTNPALDSEGSVWVTINGQDDMTIDSGFYQPKYSLGNYV 721  
Db 94 TYPYFELP-----SELSSILSNPNIRANWKID---YKIAYLGGIGTG 133  
QY 722 WYDTNKGIOG-----DDEKGISGVKVT--LKDENGNIISTTTDE----- 760  
Db 134 LF--NQGTNGSSSNFFIDSSRNAIGAKVNHLLGVGVSSTFTLTIDLLALGVLTALPS 191  
QY 761 --NGKYQF-----DNLSNGNYVHFDFKPSGMTQTTTSDG-DDEQADGEEVHV 806  
Db 192 ANDGKLDFAARTGDLGLVDLLNSN-----AARGFITTDVGADADADADADADA 242







[illegible]

Search completed: October 5, 2004, 20:53:24  
Job time : 56.7921 secs

Db 3377 ENFONNNEGTL 3387

RESULT 15

T17451

C:Species: Streptococcus parasanguinis

C:Date: 02-Sep-2000 #sequence\_revision 02-Sep-2000 #text\_change 02-Sep-2000

C:Accession: T17451

R:Wu, H.; Fives-Taylor, P.

submitted to the EMBL Data Library, October 1998

A:Description: Nucleotide sequences of the fap1 locus.

A:Reference number: Z18788

A:Accession: T17451

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-2570 <WUH>

A:Cross-references: EMBL:AF100426; NID:g3929311; PID:g3929312; PIDN:AAC79868.1

A:Experimental source: strain FW213

C:Genetics:

A:Gene: fap1

Query Match 9.7%; Score 549; DB 2; Length 2570;

Best Local Similarity 24.5%; Pred. No. 1.le-13;

Matches 254; Conservative 209; Mismatches 384; Indels 190; Gaps 32;

Qy 71 SNDQSBEEKNDVNNQNSINTDNNQIIKKEETNNYDGIKRSEDRTSTNNVDEAT 130

Db 63 SGNVFADE---TVLAKETLITTTDANEV--KLSSENF-----SEKAEEKISQSSES- 110

Qy 131 FLQKTPQDNTHLTTEEYKSESSVSSNSSIDTAQPSHTTINREESVQTSNDVSDSHVSD 190

Db 111 -----ASESVSESIS--ESVSESVTSSESVS--SESV--SESISES--VSE 151

Qy 191 PANSKIKESNTESKEENTIIQPNKKVEDSTTS-----QPSGYNTIDEKISNQDELLNL 244

Db 152 SISESISESVSESTSTSVLSSEGAASGNKATSGTKEEKQPSVRENLDKMIS--EAEVLN- 209

Qy 245 PINEYENKARPLSTTSAPSIKRVTVNQLAAEQSNVNHLLK--VTDQSIETEGYDDSEGV 302

Db 210 -----DMAARKLIITLDAEQLE-----LWKSIVATQSLEA----- 240

Qy 303 IKAHAENLIYDVTFFVEDDKVKSGDWTVDIDKNTVPSDLTDSFTPIKDN--SGEIIA- 360

Db 241 -----TKNLIGDPNATVAD-----LQIAYTTLGN--IQALGNELIKLPNGQIYAV 285

Qy 361 -----TGYDNKKQIYTTFTDYVDYKYNENKAHUKLTSYI-----DKSKV 400

Db 286 LNNTEASRAAFLRSTTTGT-----KTTFTISDFSNGG-----TQYTWAGGNANNL 330

Qy 401 PNNTKLDVEKTKALSSVNKTIITVEYQRPENRNTANLQSMET---NIDTKHNTVBQIY 456

Db 331 KNPSSISAVYDSATGKI--SWTVEYDPTTIKSPALKTKTYGIYIDTSDSKLSPT 388

Qy 457 INPLRYSAKETNVAISNGDSEGSTIIDSTIIKVKY-----VGNQMLPD----- 501

Db 389 NVLIDGAATNPTVTFYNGSGKGI EVVSGTKTGKVTKHTITPTAFSGRANDLAULEIKWL 448

Qy 502 -SNRIYDYSEYEDVTNDDYAGLNNNDVNINFGNIDSPYIIKVISKYPDNKDDYTTIOOT 560

Db 449 AATLTSPHPFEDGSKNGYGR-----YNGQTAPYVIA-----NDSGTAL--- 487

Qy 561 VTMQTTINEYTGERTASYDNTIAPSTSSGQGGDLPEKTYK--IGDVVVEDVDKDGION 619

Db 488 -----GGYQSGVNADSIPTSDTTSSESASKSESTSKSISESVISESVISGS 536

Qy 620 TNDNEKPLSNVLVTLYTPDGTSKSVRTDEDGKYQFDGLKNGLTYKITPETEGYPTPLKH 679

Db 537 VSESSESVSSESITESVSESVS-----ESISSESVSSESSESSESSESSESVSE 587

Qy 680 SGTNPALDSEGNVVMVTINGODDMTIDSGFYQTPKYSLGNTVWYDNTNKGIGQGDDEKGIS 739



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen, Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:49:18 ; Search time 128.509 Seconds  
(without alignments)  
2734.476 Million cell updates/sec

Title: US-10-806-288-15  
Perfect score: 5646  
Sequence: 1 MINKKNLLTKKKPIANKSN.....FAGLALLGKKRKNKN 1092

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5646	100.0	1092	15	US-10-378-674-8
2	4485	79.4	930	12	US-10-615-383-10
3	4485	79.4	930	16	US-10-689-082-10
4	4485	79.4	930	16	US-10-690-184-10
5	4309	76.3	892	16	US-10-661-809-21
6	4307	76.3	892	12	US-10-282-1222A-70481
7	3351.5	59.4	670	12	US-10-282-1222A-70444
8	2787	49.4	549	15	US-10-378-674-9
9	2624.5	46.5	560	15	US-10-378-674-2
10	2586	45.8	1141	12	US-10-282-1222A-70251
11	2339.5	41.4	1633	12	US-10-282-1222A-70437
12	2303.5	40.8	1742	12	US-10-615-383-4
13	2303.5	40.8	1742	16	US-10-689-082-4
14	2303.5	40.8	1742	16	US-10-690-184-4
15	2234	39.6	1385	12	US-10-282-1222A-44324

16	2205.5	39.1	1920	12	US-10-282-122A-71413	Sequence 71413, A
17	2093.5	37.1	553	12	US-10-282-122A-44457	Sequence 44457, A
18	2069	36.6	1349	9	US-09-815-242-5898	Sequence 5898, Ap
19	2069	36.6	1349	9	US-09-815-242-13137	Sequence 13137, A
20	2011	35.6	932	9	US-09-815-242-5578	Sequence 5578, Ap
21	2011	35.6	932	9	US-09-815-242-12438	Sequence 12438, A
22	1672	29.6	343	15	US-10-378-674-4	Sequence 4, Appl1
23	1634	28.9	1021	9	US-09-815-242-5471	Sequence 5471, Ap
24	1634	28.9	1021	9	US-09-815-242-12544	Sequence 12544, A
25	1629.5	28.9	936	8	US-08-781-986A-5249	Sequence 5249, Ap
26	1629.5	28.9	936	12	US-10-329-624-5249	Sequence 5249, Ap
27	1581	28.0	935	12	US-10-282-122A-44326	Sequence 44326, A
28	1562	27.7	316	15	US-10-378-674-6	Sequence 6, Appl1
29	1536	27.2	877	12	US-10-282-122A-70428	Sequence 70428, A
30	1441.5	25.5	841	9	US-09-815-242-5779	Sequence 5779, Ap
31	1441.5	25.5	841	9	US-09-815-242-12751	Sequence 12751, A
32	1121.5	19.9	513	12	US-10-282-122A-59154	Sequence 59154, A
33	1023	18.1	265	13	US-10-073-256-78	Sequence 78, Appl1
34	865	15.3	1113	9	US-09-815-242-5836	Sequence 5836, Ap
35	836	14.8	897	9	US-09-815-242-12769	Sequence 12769, A
36	782.5	13.9	903	12	US-10-282-122A-60847	Sequence 60847, A
37	708.5	12.5	1253	14	US-10-363-798-2	Sequence 2, Appl1
38	671.5	11.9	487	12	US-10-615-383-14	Sequence 14, Appl1
39	671.5	11.9	487	16	US-10-689-082-14	Sequence 14, Appl1
40	671.5	11.9	487	16	US-10-690-184-14	Sequence 14, Appl1
41	624.5	11.1	2271	12	US-10-282-122A-43924	Sequence 43924, A
42	615.5	10.9	2283	14	US-10-172-502-4	Sequence 4, Appl1
43	607	10.8	2344	9	US-09-815-242-12713	Sequence 12713, A
44	594	10.5	1831	12	US-10-282-122A-71033	Sequence 71033, A
45	554	9.8	406	16	US-10-451-467A-262	Sequence 262, App

ALIGNMENTS

RESULT 1

US-10-378-674-8  
; Sequence 8, Application US/10378674  
; Publication No. US20040006209A1  
; GENERAL INFORMATION:  
; APPLICANT: PARTI, Joseph M.  
; TITLE OF INVENTION: MONOCLONAL AND POLYCLONAL ANTIBODIES RECOGNIZING COAGULASE-NEGATI  
; FILE OF INVENTION: STAPHYLOCOCCAL PROTEINS  
; FILE REFERENCE: P07556US01/BAS  
; CURRENT APPLICATION NUMBER: US/10/378,674  
; CURRENT FILING DATE: 2003-03-05  
; PRIOR APPLICATION NUMBER: 60/361,324  
; PRIOR FILING DATE: 2002-03-05  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 8  
; TYPE: PRT  
; LENGTH: 1092  
; ORGANISM: Staphylococcus epidermidis  
; US-10-378-674-8

Query Match 100.0%; Score 5646; DB 15; Length 1092;  
Best Local Similarity 100.0%; Pred. No. 3.8e-307;  
Matches 1092; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MINKKNLLTKKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDVK	60
Db	1	MINKKNLLTKKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDVK	60
Qy	61	DSNTDDELSDNSQSSDEEKNDVINNNQSIINTDNNQIIKKEETNNYDGIKESDRTES	120
Db	61	DSNTDDELSDNSQSSDEEKNDVINNNQSIINTDNNQIIKKEETNNYDGIKESDRTES	120
Qy	121	TTNVDENEATFLQKTPQDNTHLTFEEVKGSSSVSSNSSIDTAAQPSHTTINREESVQTS	180
Db	121	TTNVDENEATFLQKTPQDNTHLTFEEVKGSSSVSSNSSIDTAAQPSHTTINREESVQTS	180
Qy	181	DNVEDSHVSDFANSKIKESNTESGKEENTIEQPNKVEDSTTSQPSGYTNIDKISNQDE	240

```
Db 181 DNVEDSHVDFANSKIKESNTSGKEENTIQPNKVKEDSTTSQSPGTYNIDKISNQDE 240
Qy 241 LNLPINEYENKARPLSTTSQAQSIKRVTVNQLAAEQSNVNHLLIKVTDQISITGEYDDSE 300
Db 241 LNLPINEYENKARPLSTTSQAQSIKRVTVNQLAAEQSNVNHLLIKVTDQISITGEYDDSE 300
Qy 301 GVIKAHAENLIYDVFVDDKVKSGDWTVDIDKNTVPSDLTSDFTPIPKIKDNGSBEIIA 360
Db 301 GVIKAHAENLIYDVFVDDKVKSGDWTVDIDKNTVPSDLTSDFTPIPKIKDNGSBEIIA 360
Qy 361 TGTVDNKKQIYTFDVTVDYKVENIKAHLLKLTSTYDKSKVPNNNTKLDVEYKLTALSSVVK 420
Db 361 TGTVDNKKQIYTFDVTVDYKVENIKAHLLKLTSTYDKSKVPNNNTKLDVEYKLTALSSVVK 420
Qy 421 TITVEYQRPENRTANLQSMFTNIDTKNHTVEQIYIINPLYSKAKETNVNINSGNDGSGST 480
Db 421 TITVEYQRPENRTANLQSMFTNIDTKNHTVEQIYIINPLYSKAKETNVNINSGNDGSGST 480
Qy 481 IIDSTIIKVKVGNQNLPSNRIRIYDSEYEDVTNDYDQALGNNDVNINFGNIDSPYI 540
Db 481 IIDSTIIKVKVGNQNLPSNRIRIYDSEYEDVTNDYDQALGNNDVNINFGNIDSPYI 540
Qy 541 IKVISKYPNKKDDYTTIQOTVTMOTTINEYTGEPRTASYDNTIAFSTSSGCGQGLDLPPEK 600
Db 541 IKVISKYPNKKDDYTTIQOTVTMOTTINEYTGEPRTASYDNTIAFSTSSGCGQGLDLPPEK 600
Qy 601 TYKIGDYVWEDVDKDGIONTNDNEKPLSNVLVTLTYPDGTSKSVRTDEBGKYQFDGLKNG 660
Db 601 TYKIGDYVWEDVDKDGIONTNDNEKPLSNVLVTLTYPDGTSKSVRTDEBGKYQFDGLKNG 660
Qy 661 LTYKIFETPGYPTLKHSGTNPALDSEGNVSWVTINGQDDMTIDSGFYQTPKYSIGNY 720
Db 661 LTYKIFETPGYPTLKHSGTNPALDSEGNVSWVTINGQDDMTIDSGFYQTPKYSIGNY 720
Qy 721 VVYDYNKDGIOGDDKEKIGSVKVTLDKENGNIISTTTTDENGKYQFDNLNSGNYIVHFDK 780
Db 721 VVYDYNKDGIOGDDKEKIGSVKVTLDKENGNIISTTTTDENGKYQFDNLNSGNYIVHFDK 780
Qy 781 PSGMTQTITDSDGDDDEQADGEEVHVITIDHDDFSDIDNGYDDSDSDSDSDSDSDSDSD 840
Db 781 PSGMTQTITDSDGDDDEQADGEEVHVITIDHDDFSDIDNGYDDSDSDSDSDSDSDSDSD 840
Qy 841 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 900
Db 841 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 900
Qy 901 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 960
Db 901 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 960
Qy 961 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1020
Db 961 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1020
Qy 1021 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1080
Db 1021 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1080
Qy 1081 LGKRRKRNKN 1092
Db 1081 LGKRRKRNKN 1092
```

## RESULT 2

```
US-10-615-383-10
; Sequence 10, Application us/10615383
; Publication No. US20040038327A1
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy
; TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLOCOCCUS
; FILE REFERENCE: P06335US03/BAS
; CURRENT APPLICATION NUMBER: US/10/615,383
```

```
; CURRENT FILING DATE: 2003-07-09
; PRIOR FILING DATE: 09/386,962
; PRIOR FILING DATE: 1999-08-31
; PRIOR FILING DATE: 1999-08-31
; PRIOR FILING DATE: 1998-08-31
; PRIOR FILING DATE: 1998-08-31
; PRIOR FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-615-383-10
```

```
Query Match 79.4%; Score 4485; DB 12; Length 930;
Best Local Similarity 81.1%; Pred. No. 2,4e-242;
Matches 883; Conservative 19; Mismatches 27; Indels 160; Gaps 3;
```

```
Qy 4 KKNLLTKKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDYKDSN 63
Db 2 KKNLLTKKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDYKDSN 61
Qy 64 TDELSDSDQSDDEKNDVINNNQSDINTDNNQIIKKEETNNYDGIKRSSEDTSTTN 123
Db 62 MDDELSDSDQSDNEEKNDVINNSQSINTDDNQ-IKKEETNSDAIENRSDITQSTTN 120
Qy 124 VDENEATFLQKTPQDNTLTEREVKSSVESNSIDTAQOPSHTTINREBSVQTSNV 183
Db 121 VDENEATFLQKTPQDNTLTEREVKSSVESNSIDTAQOPSHTTINREBSVQTSNV 180
Qy 184 EDHVSDFANSKIKESNTSGKEENTIEQPNKVKEDSTTSQSPGTYNIDKISNQDELLN 243
Db 181 ENSRVDFANSKIKESNTSGKEENTIEQPNKVKEDSTTSQSPGTYNIDKISNQDELLN 240
Qy 244 LPINEVENKARPLSTTSQAQSIKRVTVNQLAAEQSNVNHLLIKVTDQISITGEYDDSEGV 303
Db 241 LPINEVENKARPLSTTSQAQSIKRVTVNQLAAEQSNVNHLLIKVTDQISITGEYDDSEGV 300
Qy 304 KAHDAENLIYDVFVDDKVKSGDWTVDIDKNTVPSDLTSDFTPIPKIKDNGSBEIIATGT 363
Db 301 KAHDAENLIYDVFVDDKVKSGDWTVDIDKNTVPSDLTSDFTPIPKIKDNGSBEIIATGT 360
Qy 364 YDNKNKQIYTFDVTVDYKVENIKAHLLKLTSTYDKSKVPNNNTKLDVEYKLTALSSVKNKIT 423
Db 361 YDNKNKQIYTFDVTVDYKVENIKAHLLKLTSTYDKSKVPNNNTKLDVEYKLTALSSVKNKIT 420
Qy 424 VEYQRPENRTANLQSMFTNIDTKNHTVEQIYIINPLYSKAKETNVNINSGNDGSGSTIID 483
Db 421 VEYQRPENRTANLQSMFTNIDTKNHTVEQIYIINPLYSKAKETNVNINSGNDGSGSTIID 480
Qy 484 DSTIIKVKVGNQNLPSNRIRIYDSEYEDVTNDYDQALGNNDVNINFGNIDSPYIIVK 543
Db 481 DSTIIKVKVGNQNLPSNRIRIYDSEYEDVTNDYDQALGNNDVNINFGNIDSPYIIVK 540
Qy 544 ISKYDPNKDDYTTIQOTVTMOTTINEYTGEPRTASYDNTIAFSTSSGCGQGLDLPPEKTYK 603
Db 541 ISKYDPNKDDYTTIQOTVTMOTTINEYTGEPRTASYDNTIAFSTSSGCGQGLDLPPEKTYK 600
Qy 604 IGDYVWEDVDKDGIONTNDNEKPLSNVLVTLTYPDGTSKSVRTDEBGKYQFDGLKNGLTY 663
Db 601 IGDYVWEDVDKDGIONTNDNEKPLSNVLVTLTYPDGTSKSVRTDEBGKYQFDGLKNGLTY 660
Qy 664 KITFETPGYPTLKHSGTNPALDSEGNVSWVTINGQDDMTIDSGFYQTPKYSIGNYVY 723
Db 661 KITFETPGYPTLKHSGTNPALDSEGNVSWVTINGQDDMTIDSGFYQTPKYSIGNYVY 720
Qy 724 DTNKDGIQGDDEKIGSVKVTLDKENGNIISTTTTDENGKYQFDNLNSGNYIVHFDKPSG 783
Db 721 DTNKDGIQGDDEKIGSVKVTLDKENGNIISTTTTDENGKYQFDNLNSGNYIVHFDKPSG 780
Qy 784 WTQTTTSDGDDDEQADGEEVHVITIDHDDFSDIDNGYDDSDSDSDSDSDSDSDSDSDSD 843
```



```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-690-184-10

Query Match      79.4%; Score 4485; DB 16; Length 930;
Best Local Similarity 81.1%; Pred. No. 2.4e-242;
Matches 883; Conservative 19; Mismatches .27; Indels 160; Gaps 3;

QY 4 KKNLLTKKPIANKSKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDQVKDSN 63
DB 2 KKNLLTKKPIANKSKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDQVKDSN 61
QY 64 TDELSDSNQSSDEEKNDVNNQSINTDDNNOI IKKEETNNYDGLKESEDETESTTN 123
DB 62 MDELSDSNQSSDEEKNDVNNQSINTDDNNOI IKKEETNSDNLAKNSKOLITQSTTN 120
QY 124 VDENEATFLOKTPQDNTLHTEEEYKSSSVSSSSSSSSSSSSSSSSSSSSSSSSSS 183
DB 121 VDENEATFLOKTPQDNTLHTEEEYKSSSVSSSSSSSSSSSSSSSSSSSSSSSSSS 180
QY 184 EDHSVDFPANSKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQDELLN 243
DB 181 ENSRVDFPANSKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQDELLN 240
QY 244 LPINEYENKARPLSTTSAQPSIKRVTYNQLAAEQGSNNVHLIKVTDQSITTEGYDDSEGI 303
DB 241 LPINEYENKARPLSTTSAQPSIKRVTYNQLAAEQGSNNVHLIKVTDQSITTEGYDDSEGI 300
QY 304 KAHAENLIYDVFPEVDKVKSGDVTMDIDKNTVPDLTDSFTPIKIDNSGEIITATGT 363
DB 301 KAHAENLIYDVFPEVDKVKSGDVTMDIDKNTVPDLTDSFTPIKIDNSGEIITATGT 360
QY 364 YDNKNGKITFTDYDVKYENIKAKHLKLTYSIDKSKVPNNNTKLDVEYKLTALSVMKTIIT 423
DB 361 YDNKNGKITFTDYDVKYENIKAKHLKLTYSIDKSKVPNNNTKLDVEYKLTALSVMKTIIT 420
QY 424 VEYORPNEANTANLOSMTNIDTKNHTVEQTIYINPLRYSAKTNVNNISGNGEGSTIID 483
DB 421 VEYORPNEANTANLOSMTNIDTKNHTVEQTIYINPLRYSAKTNVNNISGNGEGSTIID 480
QY 484 DSTIIKVKYGDQNLNLPDSNRIDYSEYEDVTNDYLAQLGNNDVNNIFGNDSPYIIKV 543
DB 481 DSTIIKVKYGDQNLNLPDSNRIDYSEYEDVTNDYLAQLGNNDVNNIFGNDSPYIIKV 540
QY 544 ISKYDPNKDDYTIIOQTVTWQTTINEYTGFRPTASVDNTIAFTSSGQGGDLPPKTYK 603
DB 541 ISKYDPNKDDYTIIOQTVTWQTTINEYTGFRPTASVDNTIAFTSSGQGGDLPPKTYK 600
QY 604 IGDYVWEDVDKGIQNTNDNEKPLSNVLVLTYPDGTSKSVRTDEGKYQFDGLKNGLTY 663
DB 601 IGDYVWEDVDKGIQNTNDNEKPLSNVLVLTYPDGTSKSVRTDEGKYQFDGLKNGLTY 660
QY 664 KITFETPEGYPTPLKHSNTPALDSEGNSSVWVTINGQDDMTIDSGFYQTPKYSLGNYVWY 723
DB 661 KITFETPEGYPTPLKHSNTPALDSEGNSSVWVTINGQDDMTIDSGFYQTPKYSLGNYVWY 720
QY 724 DTNKGIOGDEKIGISGVKVTAKDENGNIISTTTTDENGKYQFDNLNSGNYIYHFDKPSG 783
DB 721 DTNKGIOGDEKIGISGVKVTAKDENGNIISTTTTDENGKYQFDNLNSGNYIYHFDKPSG 780
QY 784 MTQTTTDSGDDDEQADAGEVHVHTIITDHDPSIDNGYDDDESDDSDSDSDSDSDSDSDS 843
DB 781 MTQTTTDSGDDDEQADAGEVHVHTIITDHDPSIDNGYDDDESDDSDSDSDSDSDSDS 833
QY 844 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 903
DB 834 ----- 833
QY 904 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 963
```

```
DB 834 ----- 833
QY 964 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1023
DB 834 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 880
QY 1024 GSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1083
DB 881 -----NSSDKNTKDKLPDTGANEDHDSKGTLLGALFAGLGALLGK 921
QY 1084 RKQRKXKN 1092
DB 922 RKQRKXKN 930

RESULT 5
US-10-661-809-21
; Sequence 21, Application US/10661809
; Publication No. US2004010191A1
; GENERAL INFORMATION:
; APPLICANT: HOOK, Magnus
; TITLE OF INVENTION: BIOINFORMATIC METHOD FOR IDENTIFYING SURFACE-ANCHORED PROTEINS FR
; FILE REFERENCE: P07741US01/BAS
; CURRENT APPLICATION NUMBER: US/10/661,809
; PRIOR FILING DATE: 2003-09-15
; PRIOR APPLICATION NUMBER: 60/410303
; PRIOR FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 892
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-661-809-21

Query Match      76.3%; Score 4309; DB 16; Length 892;
Best Local Similarity 80.5%; Pred. No. 1.5e-232;
Matches 847; Conservative 19; Mismatches 26; Indels 160; Gaps 3;

QY 41 LFLGLGHNEAKAENSVDQVKDSNTDDELSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 100
DB 1 LFLGLGHNEAKAENSVDQVKDSNTDDELSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 59
QY 101 KEETNNYDGLKESEDETESTTNVDENEATFLQKTPQDNTLHTEEEYKSSSVSSSSSSSI 160
DB 60 KEETNSDNLAKNSKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQDELLN 119
QY 161 DTAQPSHTTINREBSVQTSNVDSHVSDPANSKIKESNTESGKEENTIEQPNKVKEDS 220
DB 120 DTAQPSHTTINREBSVQTSNVDSHVSDPANSKIKESNTESGKEENTIEQPNKVKEDS 179
QY 221 TTSQPSGYTNIDKISNQDELLNLPINEYENKARPLSTTSAQPSIKRVTYNQLAAEQGSN 280
DB 180 TTSQPSGYTNIDKISNQDELLNLPINEYENKARPLSTTSAQPSIKRVTYNQLAAEQGSN 239
QY 281 VNHLLIKVTDQSITTEGYDDSEGI KAHAENLIYDVTFEVDKVKSGDVTMDIDKNTVPS 340
DB 240 VNHLLIKVTDQSITTEGYDDSEGI KAHAENLIYDVTFEVDKVKSGDVTMDIDKNTVPS 299
QY 341 DLTDSTFTPIKIDNSGEIITATGYTNKKNQIYVTFDVTYDVKYENIKAKHLKLTYSIDKSKV 400
DB 300 DLTDSTFTPIKIDNSGEIITATGYTNKKNQIYVTFDVTYDVKYENIKAKHLKLTYSIDKSKV 359
QY 401 PNNNTKLDVEYKLTALSVMKTIITVEYORPNEANTANLOSMTNIDTKNHTVEQTIYINPL 460
DB 360 PNNNTKLDVEYKLTALSVMKTIITVEYORPNEANTANLOSMTNIDTKNHTVEQTIYINPL 419
QY 461 RYSAKETNNVNSGNGEGSTIIDSTIIKVKYVGNQNLNLPDSNRIDYSEYEDVTNDY 520
DB 420 RYSAKETNNVNSGNGEGSTIIDSTIIKVKYVGNQNLNLPDSNRIDYSEYEDVTNDY 479
QY 521 QLGNNDVNNIFGNDSPYIIKISKYDPNKDDYTIIOQTVTWQTTINEYTGFRPTASD 580
```



```
Db 480 OLGNNDVNFNIDSPYIIKISKYDPNKDYTTIQVTVMTQTTINEYTGFRFASD 539
Qy 581 NTAFSTSSGQGGDLPPPKTYKIGDYVWEDVDKQIGQNTNDNEKPLSNVLTLTPDGT 640
Db 540 NTAFSTSSGQGGDLPPPKTYKIGDYVWEDVDKQIGQNTNDNEKPLSNVLTLTPDGT 599
Qy 641 SKSVRTDEGKYQFDGLKNGLTYYKTFETPEGYTPTLKHSGTNPALDSEGNVWVINGQ 700
Db 600 SKSVRTDEGKYQFDGLKNGLTYYKTFETPEGYTPTLKHSGTNPALDSEGNVWVINGQ 659
Qy 701 DMTIDSGFYQTPKYSIGNVVYDTNKGIGQDDEKIGSVKVLKDENGNIISTTTTDE 760
Db 660 DMTIDSGFYQTPKYSIGNVVYDTNKGIGQDDEKIGSVKVLKDENGNIISTTTTDE 719
Qy 761 NGKYQFDNLNSGNYIVHFDKPSGWTQTTTDSGDDDEQADGEEVHVITTDHDDFSIDNGY 820
Db 720 NGKYQFDNLNSGNYIVHFDKPSGWTQTTTDSGDDDEQADGEEVHVITTDHDDFSIDNGY 779
Qy 821 YDESQSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 880
Db 780 YDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 795
Qy 881 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 940
Db 796 ----- 795
Qy 941 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1000
Db 796 -----DSDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 832
Qy 1001 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1060
Db 833 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 860
Qy 1061 DYCSKGTLLTFLGALIGKRRKRNKN 1092
Db 861 DHDSKGTLLGALFAGLIGKRRKRNKN 892

RESULT 6
US-10-282-122A-70481
; Sequence 70481, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Habelbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
```

```
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70481
; LENGTH: 892
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-10-282-122A-70481

Query Match 76.3%; Score 4307; DB 12; Length 892;
Best Local Similarity 80.4%; Pred. No. 2e-232;
Matches 846; Conservative 20; Mismatches 26; Indels 160; Gaps 3;

Qy 41 LFLGLGHNEAKAENSVDYKDSNTDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 100
Db 1 MFLGLGHNEAKAENTVDYKDSNMDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 59

Qy 101 KEETNYDIEKRSERDTESTTNVDENEATFLOKTPQDNTHLTTEEVKSSSVESNSSI 160
Db 60 KEETNSDAIENRSKDITQSTTNVDENEATFLOKTPQDNTHLTTEEVKSSSVESNSSM 119

Qy 161 DTAQOPSHTTINREESVQTSNVDSDSHVDSFANSKIKESNTESGKENTIEIQPNKVKEDS 220
Db 120 DTAQOPSHTTINSEASIQTSNVDSDSHVDSFANSKIIESNTESKENTIEIQPNKVREDS 179

Qy 221 TTSQPSGYTNIDKISNQDELLNLPINEYENKARPLSTTSAQPSIKRVTVNQLAAEQSN 280
Db 180 ITSQPSYKNIDEKISNQDELLNLPINEYENKVRPLSTTSAQPSKRVTVNQLAAEQSN 239

Qy 281 VNHLIKVTQOSTEGYDDSEGVKAHDAENLIYDVTFEVDDKVGSGDTMTVDIKNTVPS 340
Db 240 VNHLIKVTQOSTEGYDDSGIKAHDAENLIYDVTFEVDDKVGSGDTMTVNDIKNTVPS 299

Qy 341 DLTDSTPIKIDNSGEIATGTYDNKNKOITVTFDYDYDKYENIKAHKLKLTVIDKSKV 400
Db 300 DLTDSTFAIKIDNSGEIATGTYDNKNKOITVTFDYDYDKYENIKAHKLKLTVIDKSKV 359

Qy 401 PNNNTKLDVEYKTALSSVNTKITVEYQRPNEANTANLQSMFTNIDTKNHTVEQTIYNPL 460
Db 360 PNNNTKLDVEYKTALSSVNTKITVEYQRPNEANTANLQSMFTNIDTKNHTVEQTIYNPL 419

Qy 461 RYSAKETNVNISNGDEGSTIIDSTIIKVKYKQDNQNLPSNRIYDYSEYEDVTNDYYA 520
Db 420 RYSAKETNVNISNGDEGSTIIDSTIIKVKYKQDNQNLPSNRIYDYSEYEDVTNDYYA 479

Qy 521 QLGNNNDVNFNIDSPYIIKISKYDPNKDYTTIQVTVMTQTTINEYTGFRFASD 580
Db 480 QLGNNNDVNFNIDSPYIIKISKYDPNKDYTTIQVTVMTQTTINEYTGFRFASD 539

Qy 581 NTAFSTSSGQGGDLPPPKTYKIGDYVWEDVDKQIGQNTNDNEKPLSNVLTLTPDGT 640
Db 540 NTAFSTSSGQGGDLPPPKTYKIGDYVWEDVDKQIGQNTNDNEKPLSNVLTLTPDGT 599

Qy 641 SKSVRTDEGKYQFDGLKNGLTYYKTFETPEGYTPTLKHSGTNPALDSEGNVWVINGQ 700
Db 600 SKSVRTDEGKYQFDGLKNGLTYYKTFETPEGYTPTLKHSGTNPALDSEGNVWVINGQ 659

Qy 701 DMTIDSGFYQTPKYSIGNVVYDTNKGIGQDDEKIGSVKVLKDENGNIISTTTTDE 760
Db 660 DMTIDSGFYQTPKYSIGNVVYDTNKGIGQDDEKIGSVKVLKDENGNIISTTTTDE 719

Qy 761 NGKYQFDNLNSGNYIVHFDKPSGWTQTTTDSGDDDEQADGEEVHVITTDHDDFSIDNGY 820
Db 720 NGKYQFDNLNSGNYIVHFDKPSGWTQTTTDSGDDDEQADGEEVHVITTDHDDFSIDNGY 779
```



Best Local Similarity 98.78; Pred. No. 9.6e-148;  
Matches 542; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 52 EENSVDKSDNTDDELSDNDQSSDEEKNDVINNNQSIINTDNNQI1KKEETNNYDIE 111  
DB 1 EENSVDKSDNTDDELSDNDQSSDEEKNDVINNNQSIINTDNNQI1KKEETNNNDGIE 60  
QY 112 KRSDRTSTNNVDENATFLOKTPQDNTHLTREEVKESSESSVSIDTAQPSHTTI 171  
DB 61 KSSDRSTSTNNVDENATFLOKSPQDNTHLTREEVKESSESSVSIDTAQPSHTTI 120  
QY 172 NREESVOTSDNVDSDHVSDFPANSKIKESNTESGKEENTIEQPNKVKEDSTTSOPSGVTNI 231  
DB 121 NREESVOTSDNVDSDHVSDFPANSKIKESNTESGKEENTIEQPNKVKEDSTTSOPSGVTNI 180  
QY 232 DEKISNODELLNLPINEYENKARPLSTSAQPSIKRVTVNQLAAEQGSNNVHILIKVTDQS 291  
DB 181 DEKISNODELLNLPINEYENKARPLSTSAQPSIKRVTVNQLAAEQGSNNVHILIKVTDQS 240  
QY 292 ITGEGYDSDGVIKAHDAENLIYDVTFEVDDKVKSGDVTMTVDIDKNTVPSDLTDSFTIPKI 351  
DB 241 ITGEGYDSDGVIKAHDAENLIYDVTFEVDDKVKSGDVTMTVDIDKNTVPSDLTDSFTIPKI 300  
QY 352 KDSNGEIIATGTYDNKNKQIITYFTDVVDKYENIKAHKLKTSYIDKSKVPNNNTKLDVEY 411  
DB 301 KDSNGEIIATGTYDNKNKQIITYFTDVVDKYENIKAHKLKTSYIDKSKVPNNNTKLDVEY 360  
QY 412 KTLSSVKNKTIITVEYQRPENRNTANLOSMTNIDTKNHTVEQTIYINPLRYSAKETNNVI 471  
DB 361 KTLSSVKNKTIITVEYQRPENRNTANLOSMTNIDTKNHTVEQTIYINPLRYSAKETNNVI 420  
QY 472 SGNDEGSTIIDSTIIKVKYKVDGNQLPDSNRIDYSEYEDVTNDDYAOLGNNNDVIN 531  
DB 421 SGNDEGSTIIDSTIIKVKYKVDGNQLPDSNRIDYSEYEDVTNDDYAOLGNNNDVIN 480  
QY 532 FGNIDSPYIIKVKISKYDPNKDDYTTIOQVTMTTINEYTGEPRTASYDNTIAFSTSSGQ 591  
DB 481 FGNIDSPYIIKVKISKYDPNKDDYTTIOQVTMTTINEYTGEPRTASYDNTIAFSTSSGQ 540  
QY 592 QGGDLPPPEK 600  
DB 541 QGGDLPPPEK 549

RESULT 9  
US-10-378-674-2  
; Sequence 2, Application US/10378674  
; Publication No. US20040006209A1  
; GENERAL INFORMATION:  
; APPLICANT: PATTI, Joseph M.  
; TITLE OF INVENTION: MONOCLONAL AND POLYCLONAL ANTIBODIES RECOGNIZING COAGULASE-NEGAT  
; FILE REFERENCE: P07556US01/BAS  
; CURRENT APPLICATION NUMBER: US/10/378,674  
; CURRENT FILING DATE: 2003-03-05  
; PRIOR FILING DATE: 2002-03-05  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 560  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-10-378-674-2

Query Match 46.58; Score 2624.5; DB 15; Length 560;  
Best Local Similarity 92.34; Pred. No. 1.2e-138;  
Matches 512; Conservative 17; Mismatches 25; Indels 1; Gaps 1;

QY 46 HNEAKAESENSVDKSDNTDDELSDNDQSSDEEKNDVINNNQSIINTDNNQI1KKEETN 105  
DB 7 HHHGSEENTVDKSDNNDDLSNDQSSNEEKNDVINNNQSIINTDNNQ-1KKEETN 65

QY 106 NYDGIKRSRSDRTSTNNVDENATFLOKTPQDNTHLTREEVKESSESSVSIDTAQ 165  
DB 66 SDAIENRNSKDIQTSTNNVDENATFLOKTPQDNTHLTREEVKESSESSVSIDTAQ 125  
QY 166 PSHTTINREESVOTSDNVDSDHVSDFPANSKIKESNTESGKEENTIEQPNKVKEDSTTSOP 225  
DB 126 PSHTTINSEASIQTSDBEENSRSVDFPANSKILIESNTESNKEENTIEQPNKVEDSITSOP 185  
QY 226 SGYTNIDKISNODELLNLPINEYENKARPLSTSAQPSIKRVTVNQLAAEQGSNNVHIL 285  
DB 186 SSKYKNIDKISNODELLNLPINEYENKARPLSTSAQPSIKRVTVNQLAAEQGSNNVHIL 245  
QY 286 KVTDOQITGEGYDSDGVIKAHDAENLIYDVTFEVDDKVKSGDVTMTVDIDKNTVPSDLTDS 345  
DB 246 KVTDOQITGEGYDSDGVIKAHDAENLIYDVTFEVDDKVKSGDVTMTVDIDKNTVPSDLTDS 305  
QY 346 FTTPKIKDSNGEIIATGTYDNKNKQIITYFTDVVDKYENIKAHKLKTSYIDKSKVPNNNT 405  
DB 306 FAIPKIKDSNGEIIATGTYDNKNKQIITYFTDVVDKYENIKAHKLKTSYIDKSKVPNNNT 365  
QY 406 KLDVEYKTLSSVKNKTIITVEYQRPENRNTANLOSMTNIDTKNHTVEQTIYINPLRYSAK 465  
DB 366 KLDVEYKTLSSVKNKTIITVEYQRPENRNTANLOSMTNIDTKNHTVEQTIYINPLRYSAK 425  
QY 466 ETNVNISGNDEGSTIIDSTIIKVKYKVDGNQLPDSNRIDYSEYEDVTNDDYAOLGNN 525  
DB 426 ETNVNISGNDEGSTIIDSTIIKVKYKVDGNQLPDSNRIDYSEYEDVTNDDYAOLGNN 485  
QY 526 NDVINFGNIDSPYIIKVKISKYDPNKDDYTTIOQVTMTTINEYTGEPRTASYDNTIAF 585  
DB 486 NDVINFGNIDSPYIIKVKISKYDPNKDDYTTIOQVTMTTINEYTGEPRTASYDNTIAF 545  
QY 586 STSSGQGGDLPPPEK 600  
DB 546 STSSGQGGDLPPPEK 560

RESULT 10  
US-10-282-122A-70251  
; Sequence 70251, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931





US-10-615-383-4

Query Match 40.8%; Score 2303.5; DB 12; Length 1742;  
Best Local Similarity 34.6%; Pred. No. 4.2e-120;  
Matches 605; Conservative 146; Mismatches 313; Indels 685; Gaps 46;

QY 2 INKQNLTKKPTANKSKYAIRKFTVGTASIVIGATLLFGLGHNKAKAEN-----S 55  
DB 18 INKRVDFL-----SNKVKYSIRKFTVGTASILVATLMFGAADNEAKAEDNQLESAS 71

QY 56 VQDVKDS--NTDDELS--DSNDQSDDEKDVINN-----NQSN 91  
DB 72 KEOGGRDNENSKLQVLDLNGSHSEKTTNNVATEVKKVEAPPTSDVSKPANEAVV 131

QY 92 TDD-----NNQII---KBEETNNYDGEK-----RSEDRTESTNVNDE 127  
DB 132 TNESTKPKTTEAPTNEESIAETPKTTTQODSTKKNPSIKONLNSSTTSKESKTEH 191

QY 128 EATFLO--KTPQDNTHLTEREVK--ESSVESNSSIDTAQOPS-----167  
DB 192 STKQAMSTNKSNDLTDNDSPTQSEKTSQANNDSTNQSAFQKSPSEQKVKYTKFN 251

QY 168 -----HTTIN--REBSVOTSDNVEDSHVSDFANSKIESNTESGKEENTIPQPNKVE 218  
DB 252 DEPTQDVEHTTKLTSVSTSSVNDK--QDYTRSAV-----ASLGVDSNETAITNAVR 305

QY 219 DSTTSQSPSGYTNIDEKI-----SNODELLMLP--INEYENKARPLSTTSAQPSIKR 267  
DB 306 DNLDLKAASREQINEAIIAEALKKDFNSPDYGVDTPLALNRSQSKNSP--HKSASP---R 360

QY 268 VTNVQLAAB--QGSNVHLIKVTDQ--SITEGVDDSEGVKAHDAENLIYDVTFEVDVKV 323  
DB 361 MNLMSLAAPENSGKVNVDKVTIPTLSLKNSSNHNANNVIWPTSEQFNLKANYELDDSI 420

QY 324 KSGBTMTVDIDKNTVPSDLTDSFTPIPKIKONSGEIIATGYDNKNKQITVTFDYVDKYE 393  
DB 421 KEGDTFIKQYIRPGLELPAIKTQIRKSDGSIVANGYDKTNTTYTFTNYVDQYQ 480

QY 384 NIKAKHLKLTYSIDKSPVNNNTKLDVEYKLTALSSVKNKTIIVEXORPNEKNTANLQSNFTN 443  
DB 481 NITGSPDLIATPKRETAIKONQYPMVEVTIANEVVKDFIVDYGNKKDNTTT---AAVAN 537

QY 444 IDTKNHTVEQTIINPLR-----YSAKETN-----468  
DB 538 VDNVNNKHNVVILNQNQNPKYAKYFTVKGGEFIPGEVKVYEVTDNAMVDSFNPDLN 597

QY 469 -----VNISGNGD-----478  
DB 598 SSVKDVTSQFAPKVSADGTRVDINFARSWANGKKYIVTQAVRPTGTGVVYTWLTRDG 657

QY 479 STIIDD-----STII-----KVYKVGD-----NON-----498  
DB 658 TTNNDPVRGKSTTVTYLNGSSTAQDNFTYSLGDYVWLDKNKGVDQDDDEKGLAGVYV 717

QY 499 -LPDSN-----RI-----YDYSEYEDVT-----NDYAOIGNNNDVN-----529  
DB 718 TLKDSNNRELQRTVTDQSGHYQFDNLQNGTYTVEFAIPDNYTSPANNSTNDAIDSDGER 777

QY 530 -----INFGNIDSPYI---IKVTSKY-----DPNKD-----DYTTIQ 558  
DB 778 DGRKVVVAKGTIN--NADNMTVDGTFLTPKXNVGVWEDTNKGIQDNDKEKISGVX 835

QY 559 QTV-----TWQTTINBYTGEFTASVDN---TIAFST-----SSQGGQ3-DLPPE 599  
DB 836 VTLKNKNGDTIGITTTDTSNGKYEFTGLENGDYTIEFETPEGYPTTKQNSGSDGKDSNGT 895

QY 600 K-----TYKIGDYVWEDVDKGIQNTNNEKPLSNVJLTLYPDG 639  
DB 896 KTTVTVKADANKTIDSGFYKPTNGLGVWEDTNKGIQ--DDSEKGISGVKVTLLKDRNG 953

QY 640 TS--KSVRTDSDGKYQFDGLKNGLTIKITFTPEGYPTLKHSGTNPALDSEGNVWVTIN 698  
DB 954 NAIQTITTDASGHYQFKLENG--SYTVEFTPSGYPTKANSQDITVDSNGITTTGIIN 1012

## RESULT 13

US-10-689-082-4

; Sequence 4, Application US/10689082

; Publication No. US20040142348A1

; GENERAL INFORMATION:

; APPLICANT: FOSTER, Timothy

; TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLOCOCCUS

; FILE REFERENCE: P063350504/BAS

; CURRENT APPLICATION NUMBER: US/10/689,082

; CURRENT FILING DATE: 2003-10-21

; PRIOR APPLICATION NUMBER: 09/386,962

; PRIOR FILING DATE: 1999-08-31

; PRIOR APPLICATION NUMBER: 60/098,443

; PRIOR FILING DATE: 1998-08-31

; PRIOR APPLICATION NUMBER: 60/117,119

; PRIOR FILING DATE: 1999-01-25

; NUMBER OF SEQ ID NOS: 39

; SOFTWARE: Patent version 3.1

; SEQ ID NO 4



```
; PRIOR FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1742
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-690-184-4

Query Match      40.8%; Score 2303.5; DB 16; Length 1742;
Best Local Similarity 34.6%; Pred. No. 4.2e-120;
Matches 605; Conservative 146; Mismatches 313; Indels 685; Gaps 46;

QY 2 INKNNLLTKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNKAEEN-----S 55
Db 18 INKRVDFL-----SNKNKYKIRKFTVGTASILVGATMFGAADNEAKAEDNQLSAS 71

QY 56 VQDVKDS--NWDDEL--DSNDQSDDEKNDVIN-----NOSIN 91
Db 72 KEEQGRDRNENSKLNQVLDLNGSHSEKTTNNVNADEVKKVEAPTTSVSKPKANEAV 131

QY 92 TDD-----NQI---KKEETNNYDGEK-----RSEDRTESTNVNEN 127
Db 132 TNESTKPTTAPVNEESAETPKTSTTQODSTKKNPSSLKNNLNGSSSTTSKESKTDH 191

QY 128 EATFLQ-KTPQDNLTHLFEVEYK-ESSVESGSSNSIDTAQOPS-----167
Db 192 STKQAMSTNKSNDLTNDSPQSEKTSQAANDSTNQSAQKLDSPKSEQKVKYTKFN 251

QY 168 -----HTIN-REESVQTSNDVEDSHVDFPANSIKESNTYESKEENTIIQPNKVE 218
Db 252 DEPTQDVHEHTTKLTPSVSTDSNDK--QDYTRSAV-----ASLGVDVSNTEAITNAVR 305

QY 219 DSTTSQPSGVNTNIDEKI-----SNOBELNLP--INEYENKARPLSTTSAQPSIKR 267
Db 306 DNLDLKAASREQINEAIIAEALKQDFNPDYGVDTPLALNRSQSKSP--HKASP---R 360

QY 268 VTNQLAAB--QGSNVNHLIKVTDQ--SITEGVDDSEGVIKAHDAENLIYDVTFEVDKV 323
Db 361 MNLMSLAEPNSGKNVNDKVIKNTPLSLNKSNNHANNWIWPTSEQFNLKANYELDDSI 420

QY 324 KSGDTMTVDIDKNTVPSDLTDSFTIPKIKDNGSEIATGYDNKNKQITVTFDYVDKYE 393
Db 421 KEGDTFIKQYIRPGLELPAIKTQLRSKDSIANGVYDKTNTTTFYFYVDQYQ 480

QY 384 NIKAKHLKLTYSIDKSKVPNNNTKLDVEYKLTALSSVNTKIITVEYQRPNEENTANLQSMFTN 443
Db 481 NITGSPDLIATPKRETAIKDNQNPMEVTTIANEVVKQDFIVDYGKKNKNTTT---A-AVAN 537

QY 444 IDTKNHTVEGTIYNPLR-----YSAKETN-----468
Db 538 VDNVNNKHNEVYILNQNNQNPYAKYFSTVKNGEFIPGEVKYEVTDITNAMVDSFNPDLN 597

QY 469 -----VNI-SGNGD-----478
Db 598 SSVNKDVTSQFAPKVSADGTRVDINPARSWANGKKIIVTQAVRPTGTGVNVTYEWLTRDG 657

QY 479 STIIDD-----STII-----KVKYKGD-----NON-----498
Db 658 TTNNDYFRGTSTVTYLNKNGSSTAQGDNPYSLGDIWVLDKNNKGVQDDDEKGLAGVYV 717

QY 499 -LPDSN-----RI-----YDYSEVEDVT-----NDYALQGNNDVN-----529
Db 718 TLKDSNNRELQRTVTDQSGHYQDNQNGFYTFEFAIPDNYTPSPANNSTNDAIDSDGBR 777

QY 530 -----INFGNIDSPYI---IKVISKY-----DPNKO-----DYTTIQ 558
Db 778 DGTREKVVVAKGTIN--NADNMWTDVTGFLTPKKNVGDYVWEDTNKQIQDDNEKGLSGVK 835

QY 559 QTV-----TMOQTINEYGEFTASVDN---TIAFST-----SSGQGG-DLPPE 599
Db 836 VTILKNKNGDITIGTTTDSNGKYEFTGLENGDYITIEPTEPGYTPTKONGSDEGKDSNGT 895
```

## RESULT 15

US-10-282-122A-44324

; Sequence 44324, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:44:57 ; Search time 31.9642 Seconds  
(without alignments)  
1763.712 Million cell updates/sec

Title: US-10-806-288-15  
Perfect score: 5646  
Sequence: 1 MINKNNLLTKKKPIANKSN.....FAGLGALLGKKRKNKNKN 1092

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PGTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4490	79.5	930	4	US-09-134-001C-5314
2	4485	79.4	930	4	US-09-386-962C-10
3	2698.5	47.8	1166	4	US-09-200-6508-7
4	2303.5	40.8	1742	4	US-09-386-962C-4
5	1982	35.1	930	4	US-09-200-6508-3
6	1880	33.3	1315	4	US-09-200-6508-5
7	1629.5	28.9	936	4	US-08-956-171B-5249
8	1628	28.8	933	3	US-08-233-728-2
9	1628	28.8	933	3	US-09-421-668-2
10	1613.5	28.6	918	4	US-09-200-6508-1
11	671.5	11.9	487	4	US-09-386-962C-14
12	635.5	11.3	2137	4	US-09-134-001C-4463
13	488	8.6	1027	4	US-08-956-171B-5254
14	482	8.5	466	4	US-09-134-001C-4749
15	408.5	7.2	669	4	US-09-107-532A-6532
16	377.5	6.7	1183	4	US-09-134-001C-3530
17	349	6.2	345	3	US-08-856-253-7
18	331.5	5.9	1161	4	US-09-327-536-2
19	315.5	5.6	1112	2	US-08-714-402-2
20	301.5	5.3	1231	3	US-08-904-263A-4
21	301.5	5.3	1231	4	US-09-434-123A-4
22	295	5.2	2315	4	US-09-543-681A-5434
23	292.5	5.2	886	4	US-08-956-171B-5235
24	292.5	5.2	2504	4	US-09-328-352-5821
25	286	5.1	1060	4	US-08-911-393-2
26	282	5.0	3052	2	US-08-557-122A-26
27	282	5.0	3052	4	US-09-262-666-26

28	279	4.9	3696	4	US-09-134-001C-5080	Sequence 5080, Ap
29	278	4.9	699	4	US-09-134-001C-4054	Sequence 4054, Ap
30	276.5	4.9	1085	1	US-08-431-080-28	Sequence 28, Appl
31	276.5	4.9	1085	2	US-08-938-534-28	Sequence 28, Appl
32	276.5	4.9	1085	4	US-09-345-294-28	Sequence 2, Appl
33	276	4.9	2391	2	US-08-446-855A-2	Sequence 2, Appl
34	276	4.9	2391	3	US-09-150-741-2	Sequence 2, Appl
35	271	4.8	1115	2	US-08-568-459A-2	Sequence 2, Appl
36	271	4.8	1115	2	US-08-487-826B-2	Sequence 2, Appl
37	271	4.8	1115	2	US-09-210-288-2	Sequence 2, Appl
38	271	4.8	1115	6	5198347-6	Patent No. 5198347
39	270.5	4.8	1833	4	US-08-621-944A-4	Sequence 4, Appl
40	270.5	4.8	1833	4	US-08-945-567D-4	Sequence 4, Appl
41	270.5	4.8	1992	4	US-08-621-944A-3	Sequence 3, Appl
42	270.5	4.8	1992	4	US-08-945-567D-3	Sequence 3, Appl
43	269.5	4.8	10182	4	US-09-134-001C-3159	Sequence 3159, Ap
44	269	4.8	60	4	US-09-386-962C-19	Sequence 19, Appl
45	267	4.7	2123	3	US-08-968-685A-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1  
US-09-134-001C-5314  
; Sequence 5314, Application US/09134001C  
; Patent No. #6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; PRIOR FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 5314  
; LENGTH: 930  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-5314

Query Match	79.5%;	Score 4490;	DB 4;	Length 930;
Best Local Similarity	81.2%;	Pred. No. 1.1e-241;		
Matches	884;	Conservative 19;	Mismatches 26;	Indels 160; Gaps 3;
QY	4	KKNLLTKKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDVKDSN	63	
Db	2	KKNLLTKKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDVKDSN	61	
QY	64	TDELSDSNDQSSDEKNDVINNNQSIINDDNNQIIKKEETNNYDGIKRSSEDTSTTN	123	
Db	62	MDDELSDSNDQSSNEEKNDVINNNQSIINDDNNQ- IKKEETNSNDIAENRSKDIQTSTTN	120	
QY	124	VDENEATFLKPTQDNTLHTEEVKSSSVSSNSSIDTAQPSHTTINREESVQTSQDNV	183	
Db	121	VDENEATFLKPTQDNTLHTEEVKSSSVSSNSSMDTAQPSHTTINSEASIQTSQDNE	180	
QY	184	EDSHVSDPANSKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDEKINQDELLN	243	
Db	181	ENSRVSDPANSKIIETSNESKEENTIEQPNKVKEDSTTSQPSGYTNIDEKINQDELLN	240	
QY	244	LPINEYENKARPLSTTSAQPSIKRVTVNOLABQGSNNVHLIKVTQDSITEGVDDSEGI	303	
Db	241	LPINEYENKVRPLSTTSAQPSIKRVTVNOLABQGSNNVHLIKVTQDSITEGVDDSEGI	300	
QY	304	KAHDAENLIYDVTFFEDVDKVKSGDNTVTDIKNTVPSDLTDSFTIPKIKNSGEIATGT	363	
Db	301	KAHDAENLIYDVTFFEDVDKVKSGDNTVTDIKNTVPSDLTDSFTIPKIKNSGEIATGT	360	





Query Match 40.8%; Score 2303.5; DB 4; Length 1742;  
Best Local Similarity 34.6%; Pred. No. 6.9e-120;  
Matches 605; Conservative 146; Mismatches 313; Indels 685; Gaps 46;

QY 2 INKNNLLTKKPIANKSNKVAIRKFTVTGTSIVIGATLLFGLGHNEKABEN-----S 55  
DB 18 INKRVDFL-----SNKYKTSIRKFTVTGTSIVIGATLLFGLGHNEKABEN-----S 71

QY 56 VQDVKDS--NTDDELS--DSNDQSDDEKNDVINN-----NQSN 91  
DB 72 KEEQKGRDNEKLNQVLDNGSHSEKTTNNVNAVEKVKVEAPTTSVDSVSKPANEAVV 131

QY 92 TDD-----NQII---KKEETNNYDGEK-----RSEDRTESTTNVDEN 127  
DB 132 TNESTKPKTTEAPTVEESIAETPKTSTTQODSTEKNPDLKONLNSSTTSKESKTDH 191

QY 128 EATFLO-KTPQDNTHLFEERVK-FSSSVESNSIDTAQPS----- 167  
DB 192 STKQAMSTNKSNDLTNDSPQSEKTSQANNDSTNQAPSQKOLDSKPSQKVKYTKFN 251

QY 168 -----HTTIN-REESVQTSNVEDSHVSDFANSKIKESNTESGKEENTIEOPNKVKE 218  
DB 252 DEPTQDVEHTTKLTKTSVSTDSVNDK--QDYTRSAV-----ASLGVDSETEAITNAVR 305

QY 219 DSTSQSPSGYTNIDEKI-----SNQDELLNLP--INEYENKARPLSTTSQAQSIKR 267  
DB 306 DNLDLKAASRQINEAIIAEALKKDFNPVGVDTPLALNRSQSKNSP--HKSASP---R 360

QY 268 VTNVQLAAE--QGSNNVHLIKVTDQ--SITEGYDDSEGVKAKHAENLIYDVTFEVDVKV 323  
DB 361 NMLSLAABPNKGNKNDKVKITPTLSLNKSNHANNVITPTSNEQFNKANYELDDSI 420

QY 324 KSGDTMTVDIKNTVPDLTDSFTPIKIKONSGSIIATGTVDNKNKQITVTFTDYVDKYE 383  
DB 421 KEGTFTIKYQYTRPGGLELPAIKTQLRKSGSIVANGYDKTNTTFTTFFNYVDYQY 480

QY 384 NIKAKHLKTSYDKSVNPNNTKLDVEYKALSGSVNKTITVEYORPNENRANLQSMFTN 443  
DB 481 NITGSFOLIAIPKRETAIKONQYPMVEVTIANEVVKKDFIVDYGNKKDKNTTT---AAVAN 537

QY 444 IDTKNHTVEQTIYINPLR-----YSAKETN----- 468  
DB 538 VDNVNNKHNVEVYLNNQNNQPKYAKYFTVKNGEFIPGEVKVYEVDTDNAMVDSFNPDLN 597

QY 469 -----STII-----KVKVGD-----VNISNGD-----EG 478  
DB 598 SSVNVDVTSQFAPKVSADGTRVDINPARSMANGKKYIVTQAVRPTGTGVVYTEWLTROG 657

QY 479 STIIDD-----STII-----KVKVGD-----VNISNGD-----EG 478  
DB 658 TTNNTDFYRGTKSTVTYTLNGSSSTAQGDNPYSLIGDYVWLDKKNKGVQDDDEKGLAGVYV 717

QY 499 -LPDSN-----RI-----YDYSEYEDVT-----NDYQAQGNNDVN----- 529  
DB 718 TLKDSNNRELQRTVTDQSGHYQFDNLQNGYTFEPAIPDNTYSPANNSTNDAIDSDGER 777

QY 530 -----INFGNIDSPYI---IKVISKY-----DPNKD-----DYTTIQ 558  
DB 778 DGRKVVVAKGTIN--NADNMVDTGFLTPKYNVGDYVWEDTNKGGIQDNEKGISGVK 835

QY 559 QTV-----TMOITINEYTEPRTASVDN---TIAFST-----SSQGGG-DLPPE 599  
DB 836 VTLKKNKNGDTIGTTTTDSNGKYEFTGLENGDYTFEFTPEGYTPTKQNSGDEGKDSNGT 895

QY 600 K-----TYKIGDYVWEDVDKGIQNTNDNEKPLSNVLVTLTYPDG 639  
DB 896 KTTVVVKADANKIDSIFYFTNLGDYVWEDTNKGGIQ--DSEKGISGVKVTLKDKNG 953

QY 640 TS-KSVRTDDEGKYQFGLKNGLYTKITFTPEGYTPTLKHSNTPALDSSEGNVWVTIN 698  
DB 954 NAIGTTTTDASGHYQFKGLENG-SYTFEFTPEGYTPTKANSQDITVDSNGITTTGIIN 1012

699 GODMTIDSGFYQTPKYSLGNYVYDTNKDGIQDDEKGISGVKVTLKDENGNIISTTTT 758  
1013 GADNLIDSGFYQTPKYSGDYVWEDTNKGGIQDNEKGISGVKVTLKDENGNIISTTTT 1072

759 DENGKYQFDNLGNGYVYHFDKPSGTMQTOTTSDDDEQDADGEEVHVITITDHDDFSIDN 818  
1073 DENGKYQFDNLGNGYVYHFDKPSGTMQTOTTSDDDEQDADGEEVHVITITDHDDFSIDN 1132

819 GYDD----- 823  
1133 GYFDDSDSDSDADSDSDSDSDSDADSDADSDSDSDSDSDADSDSDSDSDSDSDSD 1192

824 ----- 823  
1193 SDSDSDADSDSDSDSDADSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1252

824 ----- 823  
1253 SDADSDADSDSDSDSDADSDSDSDSDADSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1312

824 ----- 823  
1313 SDSDSDSDSDSDSDSDADSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1372

824 ----- 823  
1373 ADSDSADSDSDSDADSD 1432

824 -----ESDSDSDSDSDSDSDS 842  
1433 SPDSDSDSDSADSD 1492

843 SPDS 902  
1493 SPDSDSADSD 1552

903 SPDS 962  
1553 SPDS 1612

963 SPDS 1022  
1613 SPDSDSDSDSDSDSDSADSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1672

1023 SGSD 1082  
1673 SPDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSKNAKDKLPDTCANEDHDSKGLTLLGTLFAGLALLG 1732

1083 KRRKNRKNK 1091  
1733 RRRK-KDNK 1740

RESULT 5  
US-09-200-650E-3  
; Sequence 3, Application US/09200650E  
; Patent No. 6680195  
; GENERAL INFORMATION:  
; APPLICANT: Patti, Joseph M.  
; APPLICANT: Foster, Timothy J.  
; APPLICANT: Hook, Magnus A.O.  
; APPLICANT: Bidhina, Deirdre Ni  
; APPLICANT: Perkins, Samuel L.  
; TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aureus  
; FILE REFERENCE: P06283US2/BAS  
; CURRENT APPLICATION NUMBER: US/09/200,650E  
; CURRENT FILING DATE: 1998-11-25  
; PRIOR APPLICATION NUMBER: 60/066,815  
; PRIOR FILING DATE: 1997-11-26  
; PRIOR APPLICATION NUMBER: 60/098,427  
; PRIOR FILING DATE: 1998-08-31  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.0

[illegible][illegible]









US-09-200-650E-1

Query Match 28.6%; Score 1613.5; DB 4; Length 918;  
Best Local Similarity 36.9%; Pred. No. 6.4e-82;  
Matches 430; Conservative 127; Mismatches 281; Indels 327;

[illegible]

Qy	47	NEAKAENSVDVKDSNTDDLSDSNDQSSDEKNDVINNNQSNINTDDNQOIKEETNN	106
Db	854	SESASTKQLSESASTSTSDSASEARKSESTKSTSLSESTSVSDSASVSTSESA	913
Qy	107	YDGLT-KRSEDRTESTTNDVDSNEATFLQKTPQDQNTLTHEEVEKSESSVESSNSIDTAQ	165
Db	914	STSVSGSTSTISDSTSTSDSASIKASESASTSKULSESVSTSDSASTSTSVSDSN	973
Qy	166	PSHTTINREBSVQTSNDVESHVSDPFANSKIKESNTESGKEENTIEQPNKVKED-STTSQ	224
Db	974	SASTSLKSTSTSVSDSTSTSTSDSASTSTSESDSASTSLSESTSTSVSDSTSTST	1032
Qy	225	PSGYTNDIKLSNQDELLNLPINEYENKAPLSTTSAPSKRVTVNQLAAEQGSNVHL	284
Db	1033	DSASMSASSESNSKS-----TSLSESTSTSLSGSTSASTDSDSASTSESDSTSTSL	1087
Qy	285	IKVTDQSI-----TEGYDDDSBGVIKAH-----DAENLIYDVTVEYDDKV-----KSGDTMT	330
Db	1088	SESTSTSLSGSTSASTDSDSASTSTSESDSTSESTSLSESTSVSDSTSESTST	1147
Qy	331	VDIDKNTVPSDLTDSFTPIPKIKNSGHIATGYDNNKKOITVFTFYDYKYENIKAHLK	390
Db	1148	SESESASASTSLSGSLSTSTISDST-----STSDSDSASTST-----	1184
Qy	391	LTSYIDKSKVPNNNTKLDVEYKTAALGSVNKTIITVEYQRPNENRTANLQSMFTNIDTKNHT	450
Db	1185	-----ESESdststslseSTstslSDststslSE-----SASTSTSESDST	1225
Qy	451	VEGTIYNPLRYSAKETNVNINSNGDEBGTIIDSTIIKVKYGDNONLPDSNRIYDSE	510
Db	1226	SESTSL-----SESTSTSVS---DSTSASTDSDASTST-SVSDSESASTS-----ISE	1269
Qy	511	YEDVINDDYAQLGNNDNVNFCNGIDSPYIIKVIKSDPNKDDYTTIIQOTVTVMQTTINEY	570
Db	1270	SLSTSVSDSTSTSTSDSASTSTSESDS-----TSESTSLSESTIS--TSVSDS	1314
Qy	571	TGEFRTASVDNTIAFTSSGQGGDLPPEKTKYKIGDYVMEDVDKGIQNTDNDNEKPLSNV	630
Db	1315	T-----SASTSDSASTSTSESES-----DSASTSLSGSTSTSLSDS	1350
Qy	631	LVTLTVPDGTSGVRDDEGKQYFDGLKNGLTKYITFETPEGVTPLKHSIGTNPALDSEG	690
Db	1351	TSTST--SDSASTST--SESSEASTSLSGSTSTSLSDSTSTSTSDSASTSTSV--SDS	1404
Qy	691	NSVWVTINGODDWTIDSGFYQTPKYSLNGYVYDNTKDGIDGDEKIGISGVKVTLKDENG	750
Db	1405	NSASTSLSG---SLSTSVSDSTSTSTSDSASTST-----ESDSEASTSLSGSTSTST	1456
Qy	751	NIISTTTTIDENGKYQFDNLNSGNYIVHFDPKPSGMTQTTTDSGDDDEQDADGEBHVHTI	810
Db	1457	DSTSTSTSD-----SASTSTSV--SESNSTSTSTSTSESLSTSVSD	1493
Qy	811	-----HDDFSIDNGYYDDESDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD	864
Db	1494	STSTSTSDSASTSTSVSDSDSASTSSSESVESTSDSESTSTSTSDSASTSTSVSESNST	1553
Qy	865	SD	924
Db	1554	SLUGSTSTSVSDSTSTSDSASTSESDSDSASTSSSESYSTSVSTSVSTSDASTSESA	1613

```

Query Match      8.6%; Score 488; DB 4; Length 1027;
Best Local Similarity 15.6%; Pct. 2.4e-19;
Matches 232; Conservative 194; Mismatches 496; Indels 264; Gaps 41;

QY      8  LITKKKTPANKSN-KYAIRFTVTGTASIVIGATLFLFGLGHN-EAKAEENSVDVKDSNTD 65
Db      1  ILHLKGDIIIVGNLRIYGIIRKHLGASVFLGTIVVGMGODKEAAASROKTTVTYENG-- 58

```



